

FEATURES source Location/Qualifiers

1. .942
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6383428"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 24.5%; Score 762.4; DB 13; Length 942;
 Best Local Similarity 96.0%; Pred. No. 1.2e-152;
 Matches 872; Conservative 0; Mismatches 26; Indels 10; Gaps 9;

QY 1504 GACGAGGGGAGGGCGGGGAGGCCCGAGACACAGAGAGTGACCGGCTGTGACCGAGTAC 1563
 Db 1 GACGAGGGGAGGGCGGGGAGGCCCGAGACACAGAGAGTGACCGGCTGTGACCGAGTAC 60

QY 1564 GCGTTTCTGAAAGAGGCTTCTCGCACTACACAGCGCTTCTGTTTGGACTCCTCGAAC 1623
 Db 61 GCGTTTCTGAAAGAGCTTCTGCACTACACAGCGCTTCTGTTTGGACTCCTCGAAC 120

QY 1624 GAGGAGACGAGGAGCCACTCGGAGAAGAGTCTCTGCTGGAAAGTCTCGCGCACATCAAG 1683
 Db 121 GAGGAGACGAGGAGCCACTCGGAGAAGTCTCTGCTGGAAAGTCTCGCGCACATCAAG 180

QY 1684 ATGACCTGTGAGTGATCAAGCAAGACTCAGACGACGGCTCCACCTCAGCAG 1743
 Db 181 ATGACCTGTGAGTGATCAAGCAAGACTCAGACGACGGCTCCACCTCAGCAG 240

QY 1744 GGCTCCTTGGAGTCTTTCAGTGCTGTGTACAGATCCAGGAGGAGGTTTATCCAGCAG 1803
 Db 241 GGCTCCTTGGAGTCTTTCAGTGCTGTGTACAGATCCAGGAGGAGGTTTATCCAGCAG 300

QY 1804 GCCCTGAGCCACTTCCAGTGATCTGTGTACAGCAATGCTGCTCCAGATGGAGCACATG 1863
 Db 301 GCCCTGAGCCACTTCCAGTGATCTGTGTACAGCAATGCTGCTCCAGATGGAGCACATG 360

QY 1864 GTCTCCTCTGTCTGTGAAGCGCTGACGAGCGCCAGGTGTGCACTTGTATGGCGCC 1923
 Db 361 GTCTCCTCTGTCTGTGAAGCGCTGACGAGCGCCAGGTGTGCACTTGTATGGCGCC 420

QY 1924 ACCTACAGCGGAGCGGAGACCCGCGAGGTGCTCCGAGAGCGCACACGCTGTG 1983
 Db 421 ACCTACAGCGGAGCGGAGACCCGCGAGGTGCTCCGAGAGCGCACACGCTGTG 480

QY 1984 GTGAGCTCAGACACAGAGAGACCGTCTGTGTGACGCTTACAGTGAACATCTGCGACG 2043
 Db 481 GTGAGCTCAGACACAGAGAGACCGTCTGTGTGACGCTTACAGTGAACATCTGCGACG 540

QY 2044 GCCCTGTGACCAATCCAAACCTGTATAGAGTGTCTGTGTACCGAAATGCCCTGGGAGC 2103
 Db 541 GCCCTGTGACCAATCCAAACCTGTATAGAGTGTCTGTGTACCGAAATGCCCTGGGAGC 600

QY 2104 CGGGGGGTGAAGTCTGTGTCAAGACTCAGACACCCCACTGCAAACTTCAGAACCTG 2163
 Db 601 CGGGGGGTGAAGTCTGTGTCAAGACTCAGACACCCCACTGCAAACTTC- GAACCTG 657

QY 2164 AGGCTGAAGAGTGGCGGATCTCCAGCTCAGGCTCGAGGACCTCTCTGAGCTCTCAT 2223
 Db 658 AGGCTGAAGA- GTGGCGGATCTCTGAGCTCAGGCTCGAGGACCTCTCTGAGCTCTCAT 716

QY 2224 GCCAATAAGAAATTTGACAGGATGGATCTCAGT- GGCAACGGCGTTGGATT- CCAGGCA 2281

QY	1624	GAGGAGAC	CAGGAGCC	ACTTGG	AGAGTCT	TCTGTG	GAAGTCT	TCGCGC	ACATCAAG	1683					
Db	121	GAGGAGAC	CAGGAGCC	ACTTGG	AGAGTCT	TCTGTG	GAAGTCT	TCGCGC	ACATCAAG	180					
QY	1684	ATGGACCT	GTTGCAGT	CGGATC	CCAAAG	CAAGCT	CAGAGC	GAAGTCT	CCACCT	TGCAGCAG	1743				
Db	181	ATGGACCT	GTTGCAGT	CGGATC	CCAAAG	CAAGCT	CAGAGC	GAAGTCT	CCACCT	TGCAGCAG	240				
QY	1744	GGCTCCT	TGAGTCT	TTC	TTCAG	CTGCTT	GCAGAT	CCAG	GAGGAG	GAGTCTTAT	TCCAGCAG	1803			
Db	241	GGCTCCT	TGAGTCT	TTC	TTCAG	CTGCTT	GCAGAT	CCAG	GAGGAG	GAGTCTTAT	TCCAGCAG	300			
QY	1804	GCCCTGAG	CCACTT	CCAGGT	GATCGT	GGT	CAGCA	CAATTC	CCCT	CCAGAT	TGGAGC	CATG	1863		
Db	301	GCCCTGAG	CCACTT	CCAGGT	GATCGT	GGT	CAGCA	CAATTC	CCCT	CCAGAT	TGGAGC	CATG	360		
QY	1864	GTCTCCT	CTGTTCT	GTGA	AGCGCT	GCAG	GAGCG	CGCCAG	TGCT	GCAC	TGTGTAT	TGCGGC	1923		
Db	361	GTCTCCT	CTGTTCT	GTGA	AGCGCT	GCAG	GAGCG	CGCCAG	TGCT	GCAC	TGTGTAT	TGCGGC	420		
QY	1924	ACCTAC	AGCCGGA	CGGG	GAAG	ACCG	CGAG	TGCT	CCG	CAGG	ACG	CAC	CGCTGTTG	1983	
Db	421	ACCTAC	AGCCGGA	CGGG	GAAG	ACCG	CGAG	TGCT	CCG	CAGG	ACG	CAC	CGCTGTTG	480	
QY	1984	GTCCAG	CTCAG	ACCAG	GAGG	ACCGT	TCTG	CTGGA	CGCCT	CAGT	GAA	CAAT	CTGG	CAGCG	2043
Db	481	GTCCAG	CTCAG	ACCAG	GAGG	ACCGT	TCTG	CTGGA	CGCCT	CAGT	GAA	CAAT	CTGG	CAGCG	540
QY	2044	GCCTGT	GCACCA	ATCCAA	CCCTGAT	AGAG	CTGTCT	CTGT	ACCGAA	TCCCT	TGGCAGC	2103			
Db	541	GCCTGT	GCACCA	ATCCAA	CCCTGAT	AGAG	CTGTCT	CTGT	ACCGAA	TCCCT	TGGCAGC	600			
QY	2104	CGGGGGT	GAAGCT	GTCT	CTG	TCAGGA	CT	TCAG	ACCC	CAACT	TCGAA	ACTT	CAGAA	CCCT	2162
Db	601	CGGGGGT	GAAGCT	GTCT	CTG	TCAGGA	CT	TCAG	ACCC	CAACT	TCGAA	ACTT	CAGAA	CCCT	660
QY	2163	GAGGCT	GAAGAG	GTGCG	GCATCT	TCAG	CTCAG	CT	TCG	GAGG	ACCTCT	CTG	CAG	ACTTCT	2222
Db	661	GAGGCT	GAAGAG	GTGCG	GCATCT	TCAG	CTCAG	CT	TCG	GAGG	ACCTCT	CTG	CAG	ACTTCT	720
QY	2223	AGCCAA	2228												
Db	721	AGGCCA	726												

```

RESULT 3
BQ000330
LOCUS
DEFINITION
AGENCOURT_8678096 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381082
5', mRNA sequence.
BQ000330
BQ000330.1 GI:22292344
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 937)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTp
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2570 row: 1 column: 11
High quality sequence start: 5
High quality sequence stop: 502.

```

RESULT 4

```
AW673661/c
LOCUS      571 bp      mRNA      linear      EST J1-APR-2000
DEFINITION ba56e01.y3 NIH MGC_10 Homo sapiens cDNA clone IMAGE:2900528.5
            similar to TR:075434 075434 ANGIOTENSIN/VASOPRESSIN RECEPTOR
            AII/AVP. ;, mRNA sequence.
ACCESSION  AW673661
VERSION     AW673661.1 GI:7538896
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            (bases 1 to 571)
REFERENCE   1  (bases 1 to 571)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Other ESTs: ba56e01.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-x@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 426.
FEATURES             source
     1..571
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:2900568"
     /cell_line="MGC36"
     /lab_host="DH10B"
     /clone_lib="NIH MGC 10"
     /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: Not I;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
     Average insert size 1.5 kb. Library prepared by Life
     Technologies."
ORIGIN
Query Match      17.4%; Score 539.8; DB 10; Length 571;
Best Local Similarity 97.7%; Pred. No. 5.6e-105;
Matches 558; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1297 GGGTTGTGCTCTTGGCGGAGATGGGCTCTGGAATC-AGAAATCTTATTGGAGGCA 1355
DB      |||
QY 1356 GGACCTCCGGAAGCAGCGGCTAGACGGGAAGACGCTCTGCTTCTTCAACATGAACAT 1415
DB      |||
QY 511  GGACCTCCGGAAGCAGCGGCTAGACGGGAACACGCTCTGCTTCTTCAACATGAACAT 452
DB      |||
QY 1416 CTTCCAGAGGACATCACTGTGACAGGTACTACAGCTTCACTCCACTTGAGTTCCAGGA 1475
DB      |||
QY 451  CTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCACTCCACTTGAGTTCCAGGA 392
DB      |||
QY 1476 ATCTTTTCAGCTATGTACTATATCTTGAACGAGGGAGGGCGGCGGAGGCGGCGGAGCA 1535
DB      |||
QY 391  ATCTTTTCAGCTATGTACTATATCTTGAACGAGGGAGGGCGGCGGAGGCGGCGGAGCA 332
DB      |||
QY 1536 GGACGTGACCGGCTGTGACCGAGTACCGGTTTCTGAAAGAGGAGTCTTCTGGCACTAC 1595
DB      |||
QY 331  GGACGTGACCGGCTGTGACCGAGTACCGGTTTCTGAAAGAGGAGTCTTCTGGCACTAC 272
DB      |||
QY 1596 GAGCGGCTCTCTGTTGGACTCTGAACGAGGAGCCAGGAGCCACTCGAGAGAGTCT 1655
DB      |||
QY 271  GAGCGGCTCTCTGTTGGACTCTGAACGAGGAGCCAGGAGCCACTCGAGAGAGTCT 212
DB      |||

AW67278
LOCUS      664 bp      DNA      linear      GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-057D11.R, genomic survey sequence.
ACCESSION  AG067278
VERSION     AG067278.1 GI:16619080
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
            1
            Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 664)
            Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan, 305-8565, Japan
            (E-mail: chimpes@gs.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
            Tel: 81-45-503-9111, Fax: 81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB. This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI
FEATURES             Location/Qualifiers
     1..664
     /organism="Pan troglodytes"
     /mol_type="genomic DNA"
     /db_xref="taxon:9598"
     /clone="PTB-057D11.R"
     /sex="male"
     /cell_type="lymphoblast"
     /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      16.7%; Score 518.8; DB 29; Length 664;
Best Local Similarity 92.7%; Pred. No. 2e-100;
Matches 555; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 750 GCCTCTCAGGAGCTCATCCGAGTCCCGAGCGCTCTCTTTTCATCATCGAGGTTTCA 809
DB      |||
QY 49  GCATGCTGAGAGCTCATCCGAGTTTCGAGGCGCTCTTTTCATCATCGAGGTTTCA 108
DB      |||
QY 810 TGAGCTCAAGCTCTCTTTCACGATCTCTCAGGAGCCCTGGTGCCTCTGCTGGAGGAA 869
DB      |||
QY 109 TGAGCTCAAGCTCTCTTTCACGATCTCTCAGGAGCCCTGGTGCCTCTGCTGGAGGAA 168
DB      |||
```


QY 870 ACGGCCACGGAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTCCTCAGCTATC 929
 Db 169 ACGGCCACGGAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTTTTGAAGTATC 228
 QY 930 TTGTCTCATCACCACAGGCCACCGCTTTGGAGAGCTCCACCGTGTGTCGAGCACCC 989
 Db 229 TTGTCTCATCACCACAGGCCACCGCTTTGGAGAGCTCCACCGTGTGTCGAGCACCC 288
 QY 990 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGCAGAAAGGAAGTAATCTTACAGTA 1049
 Db 289 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGCAGAAAGGAAGTAATCTTACAGTA 348
 QY 1050 TTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTACGTGAGGGAACAAGAGCTCT 1109
 Db 349 TTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTACGTGAGGGAACAAGAGCTCT 408
 QY 1110 CTTCCATGCTGCTGCTCCCTGTGTGCTGGGTGGTGTGACCTGCTCCAGCAGCA 1169
 Db 409 CTTCCATGCTGCTGCTCCCTGTGTGCTGGGTGGTGTGACCTGCTCCAGCAGCA 468
 QY 1170 GCTGAGGGTGGGGGCTGTTGAGACAGAGCTCCAGGACCCAGCCCGGCTCCAGCCCA 1288
 Db 469 GCTGAGGGTGGGGGCTGTTGAGACAGAGCTCCAGGACCCAGCCCGGCTCCAGCCCA 528
 QY 1230 CTACCTGTGAGTCTGATCCAAACCCAGCCCGGCTCCAGCCCGGCTCCAGCCCA 1288
 Db 529 CTACCTGTGAGTCTGATCCAAACCCAGCCCGGCTCCAGCCCGGCTCCAGCCCA 588
 QY 1289 ACCAGAGGGTGTGCTCTTGGGGGAGATGGGCTGGAATCAGAAATCTTATT 1347
 Db 589 TGCANAGAGGGTGTGCTCTTGGGGGAGATGGGCTGGAATCAGAAATCTTATT 647

RESULT 6
 AW956628
 Locus EST368698
 DEFINITION MAGe resequences, MAGD Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW956628
 VERSION AW956628.1 GI:8146311
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.
 TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 94
 Seq primer: Reverse.

FEATURES
 source
 1..568
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGD"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 15.9%; Score 494.2; DB 10; Length 568;
 Best Local Similarity 97.4%; Pred. No. 3.5e-95;
 Matches 523; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY 1638 CCACCTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGCAATCAAGATGACCTGTGCA 1597
 Db 1 CCACCTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGCAATCAAGATGACCTGTGCA 60
 QY 1698 GTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTCCACCTGCGAGCGGCTCTCTTGGAGTT 1757
 Db 61 GTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTCCACCTGCGAGCGGCTCTCTTGGAGTT 120
 QY 1758 CTTGAGTCTGTTGATCAGATCCAGAGGAGGAGTTTATCCAGCAGGCTCCAGCAGCTT 1817
 Db 121 CTTGAGTCTGTTGATCAGATCCAGAGGAGGAGTTTATCCAGCAGGCTCCAGCAGCTT 180
 QY 1818 CCAGGTGATGCTGGTCCAGCAAAATTCCTCCAGAGTGGAGCAGATGCTCTCTGCTCTG 1877
 Db 181 CCAGGTGATGCTGGTCCAGCAAAATTCCTCCAGAGTGGAGCAGATGCTCTCTGCTCTG 240
 QY 1878 TCTGAAGCGCTCAGAGGCGCCAGGTGCTGCACTTGTATGGCGCCACCTACAGCGCGA 1937
 Db 241 TCTGAAGCGCTCAGAGGCGCCAGGTGCTGCACTTGTATGGCGCCACCTACAGCGCGA 300
 QY 1938 CGGGAGAGCCCGCGAGGTGCTCCAGAGGCGCACGCTGTTGGTGGAGCTCAGACC 1997
 Db 301 CGGGAGAGCCCGCGAGGTGCTCCAGAGGCGCACGCTGTTGGTGGAGCTCAGACC 357
 QY 1998 AGAGAGGACCGTCTGCTGACGCGCTTACAGTGAATCTGGCAGCGCGCTGTGACCAA 2057
 Db 358 AGAGAGGACCGTCTGCTGACGCGCTTACAGTGAATCTGGCAGCGCGCTGTGACCAA 417
 QY 2058 TCCAAACCTGATAGAGCTGCTCTGTACCGAAATGCCCTGGGCGAGCGGGGGTGAAGCT 2117
 Db 418 TCCAAACCTGATAGAGCTGCTCTGTACCGAAATGCCCTGGGCGAGCGGGGGTGAAGCT 477
 QY 2118 GCTCTGTCAGAGACTCAGACACCCCACTGCAAACTTCAGAACCTGAGGCTGAAGAG 2174
 Db 478 GCTCTGTCAGAGACTCAGACACCCCACTGCAAACTTCAGAACCTGAGGCTGAAGAG 533

RESULT 7
 AW956628
 Locus HS_5408_A2_C04_SP6E_RPCI-11
 DEFINITION Human Male BAC Library Homo sapiens genomic clone Plate=984 Col=8 Row=E, genomic survey sequence.
 ACCESSION AW956628
 VERSION AW956628.1 GI:5485912
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 496)
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 984 row: E column: 8
 Seq primer: SP6

QY 1960 TCCGAGGAGCGCAC 1976
 Db 481 TCTCGATGCGCAC 497

RESULT 9
 BI911853
 LOCUS
 DEFINITION 764 bp mRNA linear EST 16-OCT-2001
 603063634F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212737 5',
 mRNA sequence.
 ACCESSION BI911853
 VERSION BI911853.1 GI:16175736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 764)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11534 row: c column: 10
 High quality sequence start: 22
 High quality sequence stop: 764.
 Location/Qualifiers
 1..764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5212737"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."

ORIGIN
 Query Match 13.7%; Score 426; DB 12; Length 764;
 Best Local Similarity 86.7%; Pred. No. 1.9e-80;
 Matches 539; Conservative 0; Mismatches 0; Indels 83; Gaps 3;

QY 1 ATGTACGAACCGGAGGAGCGGCTGTGGCTGTCCACTTCTGAGAACTC 60
 Db 143 ATGTACGAACCGGAGGAGCGGCTGTGGCTGTCCACTTCTGAGAACTC 202

QY 61 GAGGCTGTGGAATGAAGATTCAAGTTATATCTGGGACCGGACGAGCTGGGAGAA 120
 Db 203 GAGGCTGTGGAATGAAGATTCAAGTTATATCTGGGACCGGACGAGCTGGGAGAA 262

QY 121 GCGAAGATCCCTGGGAGAGATGGAAGAGCGGCTCCCTTGAATGCGCCAGCTGTCTC 180
 Db 263 GCGAAGATCCCTGGGAGAGATGGAAGAGCGGCTCCCTTGAATGCGCCAGCTGTCTC 322

QY 181 ATCACCACCTTGGGCGGAGGAGGCTGAGGTTGGCTCTCAGACCTTTGAGCGGATA 240
 Db 323 ATCACCACCTTGGGCGGAGGAGGCTGAGGTTGGCTCTCAGACCTTTGAGCGGATA 382

QY 241 AACAGGAGGACCTTGTGGGAGAGGACAGAGAGGACCTGTGTAGG----- 288
 Db 383 AACAGGAGGACCTTGTGGGAGAGGACAGAGAGGACCTGTGTAGGATACCCCACT 442

QY 289 ----- 288

Db 443 GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCTTCTGGAGTCTCTTGTGCTACT 502

QY 289 -----CATCCCAAGAACTTACAGGACTATGTCCGCAAGAAATTCGGCTCATG 339
 Db 503 CCAAGAAAGATATCCCAAGAAACCTACAGGACTATGTCCGCAAGAAATTCGGCTCATG 562

QY 340 GAAGACCGCAATGCGG-CCTAGGGAATGTCTCACTCAGCCAGGCTACACCCGGCT 398
 Db 563 GAAGACCGCAATGCGGCGCTTAGGGAATGTCTCACTCAGCCAGGCTACACCCGGCT 622

QY 399 CTGTCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGGCTAGCCCATCAAGATAGAGACCTCT 458
 Db 623 CTGTCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGGCTAGCCCATCAAGATAGAGACCTCT 682

QY 459 CCGGGGACAC-GCGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCT 517
 Db 683 CCGGGGACACTGCGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCT 742

QY 518 TTGAGCCAGAGGAGGCGCC 539
 Db 743 TTGAGCCAGAGGAGGCGCC 764

RESULT 10
 AQ443134
 LOCUS
 DEFINITION 449 bp DNA linear GSS 31-MAR-1999
 HS_5130_AI_B05_SP6E_RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=706 Col=9 Row=C, genomic survey sequence.
 ACCESSION AQ443134
 VERSION AQ443134.1 GI:4554473
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 449)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 706 row: C column: 9
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 449.
 Location/Qualifiers
 1..449
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Viallalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Jackson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skaleks,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3359)
Strausberg,R.
Direct Submission
Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:18204229.
Contact: MGC help desk
Email: gcapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lalic,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,I., Pearson,R., Shankripop,S., Thomas,P.J., Touchman,J.W., Tsurgoeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

FEATURES
source
location/Qualifiers
1..3359
 `/organism="Homo sapiens"`
 `/mol_type="mRNA"`
 `/db_xref="taxon:9606"`
 `/clone="IMAGE:4811303"`
 `/tissue_type="Placenta, choriocarcinoma"`
 `/clone_lib="NIH MGC 21"`
 `/lab_host="DH10B-R"`
 `/note="Vector: pOTB7"`

ORIGIN
Query Match 12.1%; Score 377; DB 11; Length 3359;
Best Local Similarity 48.9%; Pred.No. 1.7e-69;
Matches 1269; Conservative 0; Mismatches 1260; Indels 66; Gaps 7;

QY	513	CCTCTTTTGACGACAGGAGCGGCCCGGACCACGGCACCCTGTGTCATGCAAGCGC	572
Db	469	CGCGCTTTTGTCTCCCAAGAAACTGGGAACAAGCCACGTACAGTAGTATTTCAGGACC	528
QY	573	GGCAGGATAGGCAAGTCCATGCTGGCACACAAGGTGATGTCGATCGGCGACGGGA	632
Db	529	ACAAGGAATTGNAACACGACTCTGTGTAAGTGTATGATGGCTGTGTCGACACAA	588
QY	633	GCTCTTCCAAGCAGATTGTTGATTATCTTCTACATCAAATGCAGGAGATGAACCA	692

Db 1666 GAAGAGCTTAGGGAGCGTGGCAATCTCTCAGGACAGAGTGGATTCTTTGGCGATATTTTA 1725
 Qy 1764 CTGCTTGTACAGATCAGGAGGAGGATTTATCCAGCAGCGCTTGAGCCACTTCCAGGT 1823
 Db 1726 CTGTCTCTTTGAATGCAAGATCTGCTTTGTGAAGCAGCAGTGAACCTCTCCAG- 1784
 Qy 1824 GATCGTGTGACGACAACTTGTCTCAAGATGAGCACAATGTTCTCTCTGTTCTGTGAA 1883
 Db 1785 --AAGCTAACTTTCATATTTATGACACAGTGGACTTGTGTGTTCTGCTACTGCTTAA 1842
 Qy 1884 GCGTGTGAGGAGCGCCAGTGTGCTGCACTTTATGTCGCCCACTTACAGCGGGAGCGGA 1943
 Db 1843 ATACTGTCTCCAGCTTGTAGGAAACTCTGTTTTCGTTCCAAATGCTTTTAAAGAAAGGA 1902
 Qy 1944 AGACGCGCAGGCTGCTCCGAGGAGCGCACAGCTGTTGTGAGCTTGCACCAATCCAAA 2003
 Db 1903 TGAACACAGCTCTAGTTCGGATTAACAGCTCATCTGTTGGCATCACATCTGCTCTGTG-- 1960
 Qy 2004 GACCGTTCTGCTGGACGCTTACAGTGAACATCTGSCAGCGGCCCTGTGCAACCAATCCAAA 2063
 Db 1961 -----CTCACCACAGCGGCA 1977
 Qy 2064 CCTGATAGAGTGTCTGTACCGAATGCCCTGGGACGCGGGGTGAAGCTGCTGTG 2123
 Db 1978 CCTCAGAGAGCTCCAGGTGCGAGGACAGCACCTCAGCGAGTCGACCTTTGTGACCTGTGTG 2037
 Qy 2124 TCAAGGACTCAGACACCCCAACTGCAAACTTCAGAACTTGAGGCTGGAAGAGTGCAGCAT 2183
 Db 2038 TAACAGCTGAGGATCCAGCTGTGCTTCAGAGCTTGAATAAATACGTTTCCTT 2097
 Qy 2184 CTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTCATAGCAATAAGAAATTCACAG 2243
 Db 2098 TTCTGGCCAG--AGTGTCTCTCTTGTGAGGTGCTCTTTATCAGCAGACTTGAATA 2154
 Qy 2244 GATGATCTCAGTGGACAGCGGTGGATTCCAGCGATGATGCTGCTTTCGAGGGCT 2303
 Db 2155 CTTGAGCTTCACTCCCTCAGCAAACTCTCTGATGATGATGATGATGATGATGATGATGAT 2214
 Qy 2304 GCGGATCCAGCTCAGGCTCAGATGATTCAGTTGAGGAGTGTGAGTGGAGTGGAGTGGG 2363
 Db 2215 GAATACCCAGCAGCAGCAAGCTCAAGAGCTAGCGCTGGTAAATTTGACCTCTCACCCT 2274
 Qy 2364 GCTTGTGAGGAGTGGCTTCTGTGCTGGGACCAACCAACCAATCTGTTGAGTTGAGTGGACT 2423
 Db 2275 TGATTTGAAAGTCTTGTGCTGCTTCTAACCACCAACCAAGAGCTGAGTATCTGAATGT 2334
 Qy 2424 GACAGGAATGACTGGAGGATTTGGGCTGAGGTTACTATCCAGGAGCTGAGGACCC 2483
 Db 2335 ATCTGCAA--CCAGTTAGACACAGGCGGTGCCCTTTTGTGTGAAGCCCTGTGAGCCCT 2391
 Qy 2484 AGTCTGCAGACTACGAGCTTTTGGCTGAAGATCTGCGCCCTCACTGTGCTGCTGTGA 2543
 Db 2392 AGACAGGCTCTGTGATATCTGATGTTGCTTCTGCCACCTCAGGAGCAGTGTGCGA 2451
 Qy 2544 CGAGTGGCTCACTCTCAGTGTGAACCAAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAA 2603
 Db 2452 ATACATCTCTGAATGCTTCTGCTGAACAGAGCGTGCCTATCTAGACCTCAGTGCCAA 2511
 Qy 2604 TGAGCTGGGAGCTCTGGGCTGCTGCTGTGTGAGGCGCTCAGGACATCCACAGTGCA 2663
 Db 2512 TGTCTCTGAAGGAGCAGGAGTGAATACTCTCTGAGGCGCTTGAACATCCGAGCTGTGTG 2571
 Qy 2664 GCTCCAGACCTGCGGCTTGGGATCTGCGGCTGGGCTCTGCGGCTGTGAGGCTTTTC 2723
 Db 2572 CTGGATTCACTGTGTTTGTGAATGTTTATCACTGCTGCTGCTGTGAAGACCTGCG 2631
 Qy 2724 TGTGTGTCTCAGGCGCAACCAACTCCGAGGAGCTGAGTGTGAAGTTCAAGACCTGGG 2783
 Db 2632 CTCTGCTCTCATCAGCAATCAAACTTGAAGTCTCTGCAAAATGGGTTGCAATGAATCGG 2691
 Qy 2784 AGACTGGGCGCTGTGCTGCTGAGGCGCTGCAACATCCGCTGAGCTCCAGAA 2843
 Db 2692 AGATGTGGTGTGACGCTGTTGTGTCGGCTCTGACGCAATGAGTGTGCGCTTAGAGAT 2751

Qy 2844 ACTGTGCTGATAGCTGTGTGCTCTCAGAGCAAGGCTTGTGAGAACTTTTACTTCAACCT 2903
 Db 2752 TCTTGGGTGGAAGATGTGGGTTAAGAGCAGCAGCTGTGTAAAGATCTCGGCTCTGTCT 2811
 Qy 2904 GGGATCAACAGACACCTTGACCGACCTTACTGACCAACACCGCCCTAGGGGACACAGG 2963
 Db 2812 CACCTGAGTAAGACCTTGACGAGCTCACTGACCTTGAACACCTTGGACCAACAGG 2871
 Qy 2964 TGTCCGACTGTCTTGAAGCGGCTGAGCCATCTCTGGCTGCAAACTCCGAGTCTCTGTT 3023
 Db 2872 GGTGGTCTACTCTGTGAGGCTTGAGACACCCAGAGTGTCCCTGCGAGTGTCTGGGCT 2931
 Qy 3024 ATTTGGGATGACCTGAATAAATGACCCACAGTAGTGTGGAGCGCTTCGAGTAAACAA 3083
 Db 2932 GAGAAAACTGATTTGATGAGGAAACCCAGGCACCTCTGACGGCTGAGGAGAGAGAA 2991
 Qy 3084 ACCTTATTTGACAT 3098
 Db 2992 TCCTAACCTGACCAT 3006

RESULT 13
 BC012789
 LOCUS
 DEFINITION Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift errors.

ACCESSION BC012789
 VERSION BC012789.2 GI:33872518
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3360)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalley, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 REFERENCE 2 (bases 1 to 3360)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 PUBMED Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 TITLE NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:15215377.
 COMMENT Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>

183 ATCAACACGACCTTTGACCGACCTTTACCTGACCAACACCGCCCTAGGGACACACAGTGTG 242
2968 CGACTGCTTTGCAAGCGGCTGAGCCATCTCTGGCTGCAAACTCCGAGTCTCTGTGTTATTT 3027
243 CGACTGCTTTGCAAGCGGCTGAGCCATCTCTGGCTGCAAACTCCGAGTCTCTGTGTTATTT 302
3028 GGGATGGACCTGAATAATAATGACCCAGTGGTGGCAGCGCTTCGAGTAACAAACCT 3087
303 GGGATGGACCTGAATAATAATGACCCAGTGGTGGCAGCGCTTCGAGTAACAAACCT 362
3088 TATTTGGACATTTGGCTGCTGA 3108
363 TATTTGGACATTTGGCTGCTGA 383

Search completed: August 8, 2004, 07:04:46
Job time : 7585 secs

QY 2071 GAGCTGTCTCTGTACCGAAATGCCCTGGCGAGCCGGGGGG--TGAAGCTGTCTGTCTAAG 2128
Db 602 CTGTACTCTGTACCGAAATGCCCTGGCGAGCGGGGGGTGAAGCTGTGTCTGTCAAG 661
QY 2129 GACTCAGA 2136
Db 662 GACTCAGA 669

RESULT 15
BQ710940 983 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT 8487836 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6300641
DEFINITION 5' mRNA sequence.
ACCESSION BQ710940.1 GI:21849839
VERSION BQ710940.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 983)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2515 row: d column: 18
High quality sequence start: 9
High quality sequence stop: 512.
Location/Qualifiers
1..983
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6300641"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES
source
Query Match 11.4%; Score 354.6; DB 13; Length 983;
Best Local Similarity 97.4%; Pred. No. 4.8e-65;
Matches 371; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 2728 GTGCTCCAGGCCAACCAACCTCCGGGAGCTGGACTTGAAGTTCAACGACCTGGGAGAC 2787
Db 4 GGGATTGGGACGAGGCCAACCTCCGGGAGCTGGACTTGAAGTTT-AACGACCTGGGAGAC 62
QY 2788 TGGGGCTGTGTTGTGCTGAGGGCTGACATCCGCGCTGCAGATCCAGAAACTG 2847
Db 63 TGGGGCTGTGTTGTGCTGAGGGCTGCAACATCCCGCTGCAGATCCAGAAACTG 122
QY 2848 TGGCTGATAGCTGTGGCTTCACAGCCAGGGCTTGTGAGAACTTTACTTACCCCTGGGG 2907
Db 123 TGGCTGATAGCTGTGGCTTCACAGCCAGGGCTTGTGAGAACTTTACTTACCCCTGGGG 182
QY 2908 ATCAACACGACCTTGACCGACCTTTACCTGACCAACACGCCCTAGGGACACAGGTGTC 2967

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:44:41 ; Search time 58 Seconds
(without alignments)
5042.012 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 5472
Sequence: 1 MLRTAGRDGLCLSTYLEEL.....MTHSRLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5472	100.0	1035	5 ADE36417	Human PAA
2	5433	99.3	1061	5 AAO15590	Human PYR
3	5389	98.5	1099	5 AAO17857	Human PYR
4	4520	82.6	865	7 ADC31287	Human nov
5	3163	57.8	603	5 ADE36457	Human PAN
6	3005	54.9	582	5 ABU99119	Novel hum
7	2953.5	54.0	565	5 ABG37475	Human nuc
8	2694.5	49.2	521	6 ABU99120	Novel hum
9	2396	43.8	1034	4 AAE07514	Human PYR
10	2396	43.8	1034	6 ABU08503	Human PYR
11	2396	43.8	1034	6 ABU63315	Human PYR
12	2334	42.7	449	5 ABB77910	Amino aci
13	1653.5	30.2	719	5 ADE36452	Human PAA
14	1652	30.2	344	5 AAE21062	rno (upre
15	1457.5	26.6	994	5 AAO15593	Human PYR
16	1457.5	26.6	994	5 ADE36409	Human PAA
17	1456.5	26.6	920	5 ABP53254	Human MDD
18	1441	26.3	983	5 AAO17870	Human PYR
19	1432.5	26.2	896	5 AAO15592	Human PYR
20	1419	25.9	629	6 ABR01160	Human gen
21	1419	25.9	629	6 ADA98287	Human sec
22	1419	25.9	629	7 ADC20452	Human sec
23	1419	25.9	630	4 AAU01067	Human G-p
24	1385	25.3	1473	4 AAE06758	Human G-p
25	1381	25.2	1429	5 ABG78455	Human cas

ALIGNMENTS

RESULT 1
ADE36417
ID ADE36417 standard; protein; 1035 AA.

XX ADE36417;

XX 29-JAN-2004 (first entry)

DE Human PAA and nucleotide binding protein PAN6.

XX cytostatic; immunosuppressive; vulnary; antiinflammatory; vasotropic;
KW antiatheric; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; inflammatory disease; allergy;
KW restenosis; leukaemia; lymphoma; inflammatory disease; Crohn's disease;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; PAN6.

Homo sapiens.

US200307699-A1.

24-APR-2003.

25-SEP-2001; 2001US-00965621.

26-SEP-2000; 2000US-00671760.

26-SEP-2000; 2000US-0367367P.

(REED// REED J C.

(GODZ// GODZIK A.

(CHUZ// CHU Z.

(PAWL// PAWLOWSKI K.

(FIOR// FIORENTINO L.

(ARIZ// ARIZA M E.

(STEHL// STEHLIK C.

XX

PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
PI Stehlik C;
DR WPI: 2002-471256/50.
DR N-PSDB: ADE36416.
XX
PT Novel isolated PAAD domain containing polypeptide useful for inducing
PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
PT therapy for treating cancer.
XX
PS Claim 18; SEQ ID NO 24; 93pp; English.
XX
CC The invention describes an isolated PAAD domain containing polypeptide
CC (I) comprising 80% identity to the amino acid sequence of PAAD and
CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
CC -like protein containing a caspase recruitment domain (ASC)-2 fully
CC defined in specification, where (I) is biologically active. (I) is useful
CC for identifying a (I)-associated polypeptide, an agent altering that
CC association and agents that modulate PAAD domain mediated inhibition of
CC nuclear factor kappa B (NF-kappaB). A NB-ARC domain polypeptide is useful
CC for identifying an agent that modulates the activity of the NB-ARC domain
CC of (I). (I) or its functional fragments is useful in altering cellular or
CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
CC -Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (I) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, Crohn's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
CC binding protein PAN6.
XX
SQ Sequence 1035 AA;

Query Match 100.0%; Score 5472; DB 5; Length 1035;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRTAGDGLCRSLTYLEAEVLEAVELKFKLYLGTATLGEKGIPWGSMEKAGPLEMAQLL 60
DB 1 MLRTAGDGLCRSLTYLEAEVLEAVELKFKLYLGTATLGEKGIPWGSMEKAGPLEMAQLL 60
QY 61 ITHFGPEAWRLALSTPERINRKLWERGQREDLVRDQETRYDYVRKFRMLMEDRNARL 120
DB 61 ITHFGPEAWRLALSTPERINRKLWERGQREDLVRDQETRYDYVRKFRMLMEDRNARL 120
QY 121 GECVNLSHRYTRILLVKEHNPVQVQQLDTCGRGHARTVGHQASPIKTIETLPEPDERP 180
DB 121 GECVNLSHRYTRILLVKEHNPVQVQQLDTCGRGHARTVGHQASPIKTIETLPEPDERP 180
QY 181 EPRTVMVQGAAGIGKSMIAHKVWLWDWADGKLFQGRFDYLYFYNCRMNQSAECSCMDL 240
DB 181 EPRTVMVQGAAGIGKSMIAHKVWLWDWADGKLFQGRFDYLYFYNCRMNQSAECSCMDL 240
QY 241 IFSCWPEPSAPLQELIRVPRLLFIIDGFDLKPSPHDPGKWLCEWEKRPTELLNSL 300
DB 241 IFSCWPEPSAPLQELIRVPRLLFIIDGFDLKPSPHDPGKWLCEWEKRPTELLNSL 300
QY 301 IRKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKVFHNAEQGV 360
DB 301 IRKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKVFHNAEQGV 360
QY 361 FNVVRDNEPLFTMCFVPLVCWVCTCLQQOEGGLLRQTSRTTAVMYLLSLMQPKP 420
DB 361 FNVVRDNEPLFTMCFVPLVCWVCTCLQQOEGGLLRQTSRTTAVMYLLSLMQPKP 420
QY 421 GAPRLQPPPNQRLGCLSLAADGLWNQKILFQEDLRKHGLDGEDVSFLNWNIFQKDINCE 480

DB 421 GAPRLQPPPNQRLGCLSLAADGLWNQKILFQEDLRKHGLDGEDVSFLNWNIFQKDINCE 480
QY 481 RYYSFIHLSFOEFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLATSRFLGILL 540
DB 481 RYYSFIHLSFOEFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLATSRFLGILL 540
QY 541 NEETRSHLEKSLWKVSPHINKMDLLOWTOSKAQSDGSTLOQSLSEFFSCLYEIOEEEFIO 600
DB 541 NEETRSHLEKSLWKVSPHINKMDLLOWTOSKAQSDGSTLOQSLSEFFSCLYEIOEEEFIO 600
QY 601 QALSHFOVIVVSNIASRMEHVMVSFCLKRCRSAQVHLHYGATYSADGEDRARCAGAHTL 660
DB 601 QALSHFOVIVVSNIASRMEHVMVSFCLKRCRSAQVHLHYGATYSADGEDRARCAGAHTL 660
QY 661 LVQLRPRTVLLDAYSEHLAAALCTNPNLIELSYRNALSGRGVKLLCOGLRHPNCKLON 720
DB 661 LVQLRPRTVLLDAYSEHLAAALCTNPNLIELSYRNALSGRGVKLLCOGLRHPNCKLON 720
QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFPGVMLLCEGLRHPQCRLOMIQL 780
DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFPGVMLLCEGLRHPQCRLOMIQL 780
QY 781 RKCOLESAGCAQEMASVLGTNPHLVELDTGNALDELGLRLCOGLRHPVCRRLTLMKIC 840
DB 781 RKCOLESAGCAQEMASVLGTNPHLVELDTGNALDELGLRLCOGLRHPVCRRLTLMKIC 840
QY 841 RLTAACADELASTLSVNSQSLRELDLSNELGDLGVLVLLCEGLRHPTCKLQTLRLGICRLG 900
DB 841 RLTAACADELASTLSVNSQSLRELDLSNELGDLGVLVLLCEGLRHPTCKLQTLRLGICRLG 900
QY 901 SAACEGLSVVLOANHNRELDLSFNDLGDWGLMGLLAEGLOHPACRLKWLDSGLTAKA 960
DB 901 SAACEGLSVVLOANHNRELDLSFNDLGDWGLMGLLAEGLOHPACRLKWLDSGLTAKA 960
QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPGCKRLVFLFGMDLNKMTSHR 1020
DB 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPGCKRLVFLFGMDLNKMTSHR 1020
QY 1021 LAALRVTKPYLDIGC 1035
DB 1021 LAALRVTKPYLDIGC 1035
RESULT 2
ID AAO15590 standard; protein; 1061 AA.
XX AAO15590;
XX AC
XX DT 31-OCT-2002 (first entry)
XX DE Human PYRIN-8 protein #2.
XX KW Human; gene therapy; PYRIN; stress-related response; apoptotic response;
KW inflammatory response; inflammatory disorder; immune system disorder;
KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
KW autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
XX transcription profiling; PYRIN-8.
XX OS Homo sapiens.
XX PN WO200261049-A2.
XX PD 08-AUG-2002.
XX PF 31-JAN-2002; 2002WO-US002967.
XX PR 31-JAN-2001; 2001US-0265231P.
XX PR 10-SEP-2001; 2001US-0318645P.
XX

CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 865 AA;

Query Match 82.6%; Score 4520; DB 7; Length 865;
 Beat Local Similarity 93.7%; Pred. No. 0;
 Matches 865; Conservative 0; Mismatches 0; Indels 58; Gaps 2;

QY 113 MEDRNARIGECVNLSHRYTRILLVKEHNPQVQOQLDTRGHARTVGHQASPIKIETL 172
 Db 1 MEDRNARIGECVNLSHRYTRILLVKEHNPQVQOQLDTRGHARTVGHQASPIKIETL 60
 QY 173 FEPPERPEPRTVMQGAAGIGKSMARHKVMDWADGKLFQGRDYLYFYINCRMNQSA 232
 Db 61 FEPPERPEPRTVMQGAAGIGKSMARHKVMDWADGKLFQGRDYLYFYINCRMNQSA 120
 QY 233 TECSMQDLIFSCWPEPAPLQELIRVPERLLFIIDGDELKPSFHDPOGPNCLWEEKRP 292
 Db 121 TECSMQDLIFSCWPEPAPLQELIRVPERLLFIIDGDELKPSFHDPOGPNCLWEEKRP 180
 QY 293 TELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHVEHILGFSAEKKEYFYKYPH 352
 Db 181 TELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHVEHILGFSAEKKEYFYKYPH 240
 QY 353 NAEQAGQVFNVRDNEPLFTMCVFPLVCWVCTCLOQLEGGLLRQTSRTTAVMYLYL 412
 Db 241 NAEQAGQVFNVRDNEPLFTMCVFPLVCWVCTCLOQLEGGLLRQTSRTTAVMYLYL 300
 QY 413 LSLMQPKGAPLQPPNQRGLCSLAADGLWNQKILFEQDLRKHGLDGEDVSALNNMI 472
 Db 301 LSLMQPKGAPLQPPNQRGLCSLAADGLWNQKILFEQDLRKHGLDGEDVSALNNMI 360
 QY 473 FOKDINCERYGFHLSFOEPPAAMYIILDEGEGAGPDQDVTLLITEYAFSERSFLALT 532
 Db 361 FOKDINCERYGFHLSFOEPPAAMYIILDEGEGAGPDQDVTLLITEYAFSERSFLALT 420
 QY 533 SRFLFGLLNEETRSHLEKSLCWKVSPIHKWDLQWISKAQSDGSTLQQGSLEFFSCLYE 592
 Db 421 SRFLFGLLNEETRSHLEKSLCWKVSPIHKWDLQWISKAQSDGSTLQQGSLEFFSCLYE 480
 QY 593 IQEEFIIQALSHFQVIVVSNIAKMEHNVSFCLKRCESAQVHLHYGATYSGADGEDRAR 652
 Db 481 IQEEFIIQALSHFQVIVVSNIAKMEHNVSFCLKRCESAQVHLHYGATYSGADGEDRAR 540
 QY 653 CSAGAHTLVQLRPERTVILLDAYSEHLAAALCTNPMLIELSLYRNALSGRGVKLLCQGLR 712
 Db 541 CSAGAHTLVQLRPERTVILLDAYSEHLAAALCTNPMLIELSLYRNALSGRGVKLLCQGLR 599
 QY 713 HPNCKLQNLRLKRCRITSSACEDLSAALITANKMLTRMDLSGNGVGPPGMWLLCEGLRHQ 772
 Db 600 HPNCKLQNLRLKRCRITSSACEDLSAALITANKMLTRMDLSGNGVGPPGMWLLCEGLRHQ 659

QY 773 CRLQMIQIRKQLESAGACQEMASVLGTPNPHLVDELDTGNALEDLGLRLCCQGLRHPVCR 832
 Db 660 CRLQMIQIRKQLESAGACQEMASVLGTPNPHLVDELDTGNALEDLGLRLCCQGLRHPVCR 719
 QY 833 RTLWLKICRLTAACDELAELASTLSVNQSLRELDLSINELGDLGVILLCEGLRHPTCKLQTL 892
 Db 720 RTLWLKICRLTAACDELAELASTLSVNQSLRELDLSINELGDLGVILLCEGLRHPTCKLQTL 779
 QY 893 RLGICRLGSAACEGLSVVVLQANHNLRELDLSFNLDGMDGLWLLAELGLOHPACRLQKWL 952
 Db 780 R-----LD 782
 QY 953 SCGLTAKACENLYFTLGINQTLTDLTYLTNNALGTGVRLLCCKLSHPGCKLRVLWLFQMD 1012
 Db 783 SCGLTAKACENLYFTLGINQTLTDLTYLTNNALGTGVRLLCCKLSHPGCKLRVLWLFQMD 842
 QY 1013 LNKMTSHSLAALRVTKPYLDIGC 1035
 Db 843 LNKMTSHSLAALRVTKPYLDIGC 865
 RESULT 5
 ADE36457
 ID ADE36457 standard; protein; 603 AA.
 XX AC ADE36457;
 XX DT 29-JAN-2004 (first entry)
 XX DE Human PAN6 leucine-rich-repeat domain (LLR) seq id 64.
 KW cytostatic; immunosuppressive; vulnary; antiinflammatory; vasotropic;
 KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
 KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
 KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
 KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
 KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
 KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
 KW apoptosis; NFkappaB induction; cytokine processing;
 KW cytokine receptor signaling caspase-mediated proteolysis;
 KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
 KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
 KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
 KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
 KW arthritis; lupus; schrojen's syndrome; Crohn's disease; heart failure;
 KW ulcerative colitis; graft versus host disease; stroke; Alzheimer's disease;
 KW neurodegenerative disease; parkinson's disease; Alzheimer's disease;
 KW cancer therapy; PAAD domain family; human; PAN6; leucine-rich-repeat;
 LLR.
 XX OS Homo sapiens.
 XX PN US2003077699-A1.
 XX PD 24-APR-2003.
 XX PF 25-SEP-2001; 2001US-00965621.
 XX PR 26-SEP-2000; 2000US-00671760.
 XX PR 26-SEP-2000; 2000US-0367367P.
 XX (REED/) REED J C.
 XX (GODZ/) GODZIK A.
 XX (CHUZ/) CHU Z.
 XX (PAWL/) PAWLOWSKI K.
 XX (FIOR/) FIORENTINO L.
 XX (ARIZ/) ARIZA M E.
 XX (STEH/) STEHLIK C.
 XX Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;

DR WPI; 2002-471256/50.
 XX Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX Claim 24; SEQ ID NO 64; 93pp; English.
 XX The invention describes an isolated PAAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical processes such as apoptosis, NFkappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6 leucine-rich-repeat domain (LRR).
 XX Sequence 603 AA;
 SQ

Query Match 57.8%; Score 3163; DB 5; Length 603;
 Best Local Similarity 100.0%; Pred. No. 8.9e-294;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNORGCSLAADGLWQKILFEQDLKXGLDGEDVSAFLNMIFQKINCERYYSFIHL 498
 DB 1 PNORGCSLAADGLWQKILFEQDLKXGLDGEDVSAFLNMIFQKINCERYYSFIHL 60
 QY 489 SFOEFAAMYIIDEDEGAGPQDVTRLLTEYAFSERSFALTSRFLGLNEETRSHL 548
 DB 61 SFOEFAAMYIIDEDEGAGPQDVTRLLTEYAFSERSFALTSRFLGLNEETRSHL 120
 QY 549 EKSLCWKVSPIHKMDLQWTSKAQSDGSTLQOGLSEFFSCLYEIOEEFFIQALSHFQV 608
 DB 121 EKSLCWKVSPIHKMDLQWTSKAQSDGSTLQOGLSEFFSCLYEIOEEFFIQALSHFQV 180
 QY 609 IIVSNITASMEHVMVSFCLRCRCSAOLVHLGYATSGADGEDRARCAGAHLLVQRPER 668
 DB 181 IIVSNITASMEHVMVSFCLRCRCSAOLVHLGYATSGADGEDRARCAGAHLLVQRPER 240
 QY 669 TVLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCQGLRHPCNKLQNLRLKRCRI 728
 DB 241 TVLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCQGLRHPCNKLQNLRLKRCRI 300
 QY 729 SSSACEDLSAALITANKNLTMDLSGNGVGPFGPMMLICEGLRHPCQLQWILKPKQLESG 788
 DB 301 SSSACEDLSAALITANKNLTMDLSGNGVGPFGPMMLICEGLRHPCQLQWILKPKQLESG 360
 QY 789 ACOEMASVLGTNPHLVELDTGNALDGLRLCQGLRHPCQLRTRTLWLKICRLTAACD 848
 DB 361 ACOEMASVLGTNPHLVELDTGNALDGLRLCQGLRHPCQLRTRTLWLKICRLTAACD 420
 QY 849 ELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 908
 DB 421 ELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 480
 QY 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGHQHPACRLQKLMWLDSCGLTAKACENLYFTL 968

DB 481 VVLQANHNLRELDLSFNDLGDWGLWLLAEGHQHPACRLQKLMWLDSCGLTAKACENLYFTL 540
 QY 969 GINQOTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRVLWLFQMDLNKMTSHRLAALRVTK 1028
 DB 541 GINQOTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRVLWLFQMDLNKMTSHRLAALRVTK 600
 QY 1029 PVL 1031
 DB 601 PVL 603
 RESULT 6
 ABU99119
 ID ABU99119 standard; protein; 582 AA.
 AC ABU99119;
 XX
 DT 01-AUG-2003 (first entry)
 XX
 DE Novel human GPCR related protein NOV2a.
 XX
 KW Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytosolic; cardiac; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX
 OS Homo sapiens.
 XX
 PN WO200299116-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US017428.
 XX
 PR 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-031550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.
 PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 PI Gangalli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
 PI Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 PI Voss EZ, Zerhusen BD;
 XX WPI; 2003-140627/13.
 DR N-PSDB; ACD03623.
 XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.

PS Claim 1; Page 99; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections; multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV

XX Sequence 582 AA;

Query Match 54.9%; Score 3005; DB 6; Length 582;
Best Local Similarity 99.3%; Pred. No. 1.2e-278;
Matches 578; Conservative 0; Mismatches 0; Indels 4; Gaps 3;

QY 373 MCFVPLVWVCTCLOQLEGGGLRQTSRTTAVTMLYLLSIMQPKGAPRLQPPNOR 432
DB 1 MCFVPLVWVCTCLOQLEGGGLRQTSRTTAVTMLYLLSIMQPKGAPRLQPPNOR 60
QY 433 GLCSLAADGLWQKILFEQDRLKHLGDGDSVAFNLMMIFOKDINCERYYSFIHLSFOE 492
DB 61 GLCSLAADGLWQKILFEQDRLKHLGDGDSVAFNLMMIFOKDINCERYYSFIHLSFOE 120
QY 493 FFAAMYIILDEGGAGPQDVTRLLETFAYFSERSFLATSRFLGGLNNEETRSHEKSL 552
DB 121 FFAAMYIILDEGGAGPQDVTRLLETFAYFSERSFLATSRFLGGLNNEETRSHEKSL 180
QY 553 CWKVSPIIKMDLLOWTQSKAQSDGSTLQOQSLEFFSCFLYEIOEEFIQALSHFQIVVVS 612
DB 181 CWKVSPIIKMDLLOWTQSKAQSDGSTLQOQSLEFFSCFLYEIOEEFIQALSHFQIVVVS 240
QY 613 NTASKMEHWSSPCLAKRCSAQLVHLGYATYSADGEDRACSGAGHTLLVQLRPETVLL 672
DB 241 NTASKMEHWSSPCLAKRCSAQLVHLGYATYSADGEDRACSGAGHTLLVQLRPETVLL 300
QY 673 DAYSEHLAALCTNPMLIELSYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISS 731
DB 301 DAYSEHLAALCTNPMLIELSYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISS 360
QY 732 ACEDLSAALIANKNLFRMDLSGNGVFPGMMLICEGLRHPOCPLQMLQKRCOLESGACQ 791
DB 361 ACEDLSAALIANKNLFRMDLSGNGVFPGMMLICEGLRHPOCPLQMLQKRCOLESGACQ 420
QY 792 EMASVLGTNPMLVELDTGNALDGLRLCCOGLRHPVCRRLTLW-LKICRLTAACDE 849
DB 421 EMASVLGTNPMLVELDTGNALDGLRLCCOGLRHPVCRRLTLW-LKICRLTAACDE 480
QY 850 LASTLSVNSQLRELDLSINELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACBGLS 908
DB 481 LASTLSVNSQLRELDLSINELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACBGLS 540
QY 909 VVIQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKLM 950

Db 541 VVIQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKLM 582

RESULT 7

ABG97475

XX ABG97475 standard; protein; 565 AA.

AC ABG97475;

DT 16-DEC-2002 (first entry)

DE Human nucleic acid associated protein, NAAP9, from, INCYTE no.429930CDL.

XX Human; nucleic acid associated protein; NAAP; cancer;
KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.

OS Homo sapiens.

PN WO200272630-A2.

XX 19-SEP-2002.

XX 07-FEB-2002; 2002WO-US003844.

XX 09-FEB-2001; 2001US-0268118P.

XX 21-FEB-2001; 2001US-0270963P.

XX 22-FEB-2001; 2001US-0270858P.

XX 23-FEB-2001; 2001US-0271194P.

XX 07-MAR-2001; 2001US-0274071P.

XX 12-APR-2001; 2001US-0283496P.

XX 09-NOV-2001; 2001US-0344650P.

XX (INCY-) INCYTE GENOMICS INC.

XX Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Wallia NK, Thangavelu K;
PI Elliott VS, Marquis JP;

XX WPI; 2002-723320/78.

XX N-PSDB; ABS78719.

XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.

PS Claim 1; Page 147-149; 162pp; English.

XX The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody
CC screening for a compound that is effective as an ant/agonist or modulator
CC of NAAP, generating an array comprising different nucleotide molecules
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridisable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders

CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
 CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
 CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
 CC other diseases and disorders listed in the specification. These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of NAAp. The NAAp or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence represents an NAAp protein
 CC
 XX Sequence 565 AA;
 SQ

Query Match 54.0%; Score 2953.5; DB 5; Length 565;
 Best Local Similarity 99.8%; Pred. No. 1e-273;
 Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	470	MNIPQDINERYYSRTHLSFQFFFAAMYYILDEGGAGDQDVTLTLEYAFSERSFL	529
Db	1		60
QY	530	ALTSRFLGLLNEETSHLEKSLCWKYSPIHKMDLLQWIOSKQSDGSTLQOQSLRFFSC	589
Db	61		120
QY	590	LYIOBEETFOALSHFQVIVVNIASKMEHMYSSFCFKCRSAQVHLHYGATYSADGED	649
Db	121		180
QY	650	RARCSAGHTLLVQLPERVLLDAYSEHLAALCTNPNIETSLVENALGSGVKKLQ	709
Db	181		239
QY	710	GLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR	769
Db	240		299
QY	770	HPQCRLOMIOLKQLESAGCQEMASVLGNPHLVELDLTGNALEDGLRLLCQGLRHPV	829
Db	300		359
QY	830	CRRLTWLTKICRTAAACDELASTLSVNQSLRELDLSLNEGLDGLVLLCEGLRHPCKL	889
Db	360		419
QY	890	QTLRLGICRIGSAACGLSVVLQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKL	949
Db	420		479
QY	950	WLDSCGLTAKACENLYFTIGINTLDTLYTNALGDTGVRLLCCKRLSHPGCKLRVLNLF	1009
Db	480		539
QY	1010	GMDLNKMTSHRLAALRVTKPYLDIGC 1035	
Db	540		

RESULT 8

ID ABU99120

XX standard; protein; 521 AA.

AC ABU99120;

XX 01-AUG-2003 (first entry)

DT Novel human GPCR related protein NOV2b.

DE Human; G-protein coupled receptor related protein; GPCR related protein;

XX

KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX
 OS Homo sapiens.
 XX WO200299116-A2.
 XX
 XX 12-DEC-2002.
 PD
 PF 04-JUN-2002; 2002WO-US017428.
 XX
 XX 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.
 PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 PI Gangoli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI MacDougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;
 PI Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CM;
 PI Voss EZ, Zerhusen BD;
 XX
 DR WPI; 2003-140627/13.
 DR N-PSDB; ACD03624.
 DR
 XX
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS
 PS Claim 1; Page 100; 332pp; English.

The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of

QY 561 KMDLLQWIOSKAQSDGSLTQQSLEFFSCLYEIQEEFTQQALSHFQVIVVSNIAKMEH 620
 Db 603 RLELLKWLVEKAKAKLQIQPSQLFCLYEQBEDFVQRAWDYFPKTEI-NLSTRMDH 661
 QY 621 MVSSFLKRCRAQVHLHYGATYSDGEDRARCAGATHLLVQLRPRTVILLDAYSEHLA 680
 Db 662 MVSSFCIENHVESLSL-GFLHNPKEEBEKEGRHLDWQ----- 703
 QY 681 AALCTNPNIELIYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSSACEDLSAAL 740
 Db 704 ---CVLPS-----SSHAACSHG-----LVNSHLTSSFCRGLFSL 735
 QY 741 IANKNLTMDLSNGVGPFGWMLCEGLRHPQCRLOMIOLRKCQLESAGCQEWASVLGTN 800
 Db 736 STSQSLTELDLSDNSLGDPMRVLCETLQHPGNCIRRLMIGRGCSHECCFDISLIVSSN 795
 QY 801 PHLVDELDTGNALDELGLRLCQGLRHPVCRILRLTKLCRTAAACDLASTLSVNOSL 860
 Db 796 QKLVDELSDNALGDFGIRLLCVGLKHLNCKLWLVSSCCLTSACCCQDLASVLSHSL 855
 QY 861 RELDLSNELDGLVLLCEGLRHPTCKLQTLRLGLCRIGSAACGLSVVLQANHLREL 920
 Db 856 TRLYVGENALGDSGVAILECAKAPQCNIQKLGVLNSGLTSVCCSALSLSVLTNQNLTHL 915
 QY 921 DLSFNDLGDWGLWLLAEGLOHPACRLOKWLSDSCGLTAKACENLYFTLGINOTLTLXLT 980
 Db 916 YURGMTLGDGKHLKCEGLHDPCKLQVLELONCNITSCCWDLSLTLSQSRLKLSLG 975
 QY 981 NNALGDTGVRLLCRSLHSPGCKLRVLWLPMDLNKMTSRRLAALRVTKPYLDI 1033
 Db 976 NNDLGLDGLWMECEVLKQOSCLLQNLGLSEMYFNVTYSKSALETLOEKEPALT 1028
 RESULT 10
 ABU08503
 ID ABU08503 standard; protein; 1034 AA.
 AC ABU08503;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 DE Huma PYRIN-1 protein.
 XX
 KW Human; nucleotide binding site; pyrin domain; NBS-1; PYRIN-1; caspase-1;
 KW antiinflammatory; apoptosis; ASC; NF-KB; nuclear factor KB; IRR;
 KW leucine rich repeat; inflammatory disorder; familial cold urticaria;
 KW arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 KW ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 KW insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 KW psoriasis; graft versus host disease; food allergy; conjunctivitis;
 KW chronic obstructive pulmonary disease; food allergy;
 KW helminthic infection; leishmaniasis; viral infection; HIV infection;
 KW bacterial infection; tuberculosis; leprosy; chromosome 1q44.
 XX
 OS Homo sapiens.
 XX
 XX
 FN US2002187922-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 22-APR-2002; 2002US-00127516.
 XX
 PR 17-FEB-2000; 2000US-00506067.
 PR 01-SEP-2000; 2000US-00653901.
 PR 26-SEP-2001; 2001US-00964955.
 PR 20-DEC-2001; 2001US-00027629.
 XX
 XX (BERT/) BERTIN J.
 PA (MANJ/) MANJI G A.
 XX
 PI Bertin J, Manji GA;
 XX
 XX WPI; 2003-328763/31.
 Db

DR N-PSDB; ABX93556.
 XX
 PT Identification of compound that binds to polypeptide (for e.g. a PYRIN
 PT protein), useful for treating disorders associated with inappropriate
 PT apoptosis, for e.g. inflammatory disorder.
 XX
 PS Claim 1; Fig 4; 74pp; English.
 XX
 CC The invention relates to a new method for the identification of a
 CC compound that binds to a human PYRIN-1 (an apoptotic signalling molecule
 CC appearing as ABU08503) comprising contacting the polypeptide or a cell
 CC expressing the polypeptide to a test compound and determining whether
 CC PYRIN-1 binds to the test compound, or determining the effect of the test
 CC compound on the activity of PYRIN-1. In a similar manner, compounds are
 CC isolated which modulate the binding of PYRIN-1 to ASC (comprising testing
 CC the compounds against the pyrin binding domains of both PYRIN-1 and ASC
 CC (not defined)), modulate the activity of NF-kB (nuclear factor KB), and
 CC which modulate the ASC-mediated activation of NF-kB (comprising: (a)
 CC measuring the binding of a test compound to the LRR (leucine rich repeat)
 CC domain of PYRIN-1; and (b) measuring the activation of NF-kB in a cell
 CC expressing ASC and PYRIN-1 in the presence and absence of the LRR domain
 CC binding compound). The candidate modulators may be useful for treating an
 CC inflammatory disorder. Also included is a method (M8) for identifying a
 CC modulator of caspase-1 activity. The identified compound can be used to
 CC modulate the polypeptide's activity (ASC and NF-kB activities in a
 CC patient). By modulating the expression or activity of the polypeptide
 CC (PYRIN-1 and ASC), a disorder associated with inappropriate apoptosis
 CC (inflammatory disorders e.g. familial cold urticaria, arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple
 CC sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus
 CC host disease, asthma, chronic obstructive pulmonary disease, allergies
 CC (e.g. food allergies), conjunctivitis, helminthic infection (e.g.
 CC leishmaniasis), viral infections such as HIV infection, and bacterial
 CC infections such as tuberculosis and leprosy) can be treated. Human NBS-1
 CC protein (nucleotide binding site protein 1) contains a pyrin domain which
 CC was used to identify genomic sequences containing the PYRIN-1 gene. The
 CC gene for pyrin-1 is located on chromosome 1q44. The present sequence
 CC represents human PYRIN-1
 XX
 SQ Sequence 1034 AA;

Query Match 43.8%; Score 2396; DB 6; Length 1034;

Best Local Similarity 46.1%; Pred. No. 8.5e-220; Indels 100; Gaps 10;
 Matches 495; Conservative 175; Mismatches 303;

QY 11 CRLSTYLELEAVELEKFKLYL-GTATELGEKGIWGSMEKAGPLEMAQLLTHFGPEPA 69
 Db 6 CKLARYLEDELDVLLKKFMHLEDYPPQKGCIPLRGQTEKADHVDLATLMDFNGEKA 65
 QY 70 WRLALSTTERINRDLWEGQRED----- 93
 Db 66 WAMAYWIFAAINRRDLYEKAKDEPKWGSNDARVSNPTVICOEDSIEEEMGLLEYLSRI 125
 QY 94 ----LVDPDPOETRYDRYVRKPRLMEDRNARLCECVNLSHRYTRLLLVKHSNPMOVQOOL 149
 Db 126 STCKMKKYRKYRYKYSRFOCIEDRNARLGESVSLNKRYTRLRILKEHRSQREQL 185
 QY 150 LDTGRGHARTVGHQASPIKTIETLFEPEDEPERPPRTVVMQGAAGIKSMIAHKVMDWAD 209
 Db 186 LAIGK-TKTCSPVSPKIMELLFDPDDEHSESPVHTVVFQGAAGIKTILARKMWLDWAS 243
 QY 210 GKLFQGRFDYLYINCREMNQSAATSCMDLIFSCWPEPSAPQLBLIRVPELLIIDGF 269
 Db 244 GTLYQDRFDYLYIHCREVS-LVTQSLGDLIMSCCPDPNPPIHKIVRKPSILFMDGF 302
 QY 270 DELKPSFHDPOGPGWCLCWEEKRPTLLNLSLRKLLPELSLITTRPALKHLHLH 329
 Db 303 DELQCAFDEHIGLPLCTDQWKAERGDIILSLLRKLLPEASLITTRPALKHLHLH 362
 QY 330 PRHVEILGFSEAEKKEYFYKYFHNAEQAGQVFNVRDNEPFTMCVFPLVCWVCTCLOQ 389
 Db 363 PRHVEILGFSEAEKKEYFFKYFSDQAARAAASLIQENEVLFTMCVFPLVCWVCTGLKQ 422

QY	390	OLEGGLLQTSRTTAVYMLYLLSLMOPKAPRLQPPNQRGLCSLAADGLWNOKILF	449	FT	/label= Casein_kinase_II_phosphorylation_site
Db	423	QMSGKSLAQTSTKTTTAVYVFFLSLLQPRGSGEHLCAHLWGLCSLAADGLWNOKILF	482	FT	/note= "CAMP/cGMP-dependent protein kinase phosphorylation site"
QY	450	EBQDLRKHGDLGDEVSFAFLNMNIFQDINCERYYSFIHLSPOEFFAAMYIILDEGGGAG	509	FT	/label= Casein_kinase_II_phosphorylation_site
Db	483	EBSDLRNHGLQKADVSFAFLRMNLFQKEVDCERFYSFIHMTFQEFFAAMYIILDEEGRT	542	FT	/label= Expanded nucleotide binding site_domain
QY	510	-----PDQVTLLEYAFSPERSFPLATSRFLFGLNNEHTRHLEKSLCWKUSPHI	560	FT	/note= "NACHT NTPase"
Db	543	NVPSGLKLPFRDVTVLLNENYKFERGYLFFVVRFLGLVNOERTSYLEKKLSCKTSQOI	602	FT	/note= "Expanded nucleotide binding site domain signature motif #1"
QY	561	KMDLLOWIOSKAGDSGTQQGSLFFSCLYEIOEEFIQALSHFQVIVWSNIASKMHH	620	FT	/label= Kinase_1a_domain
Db	603	RUELLKWEIVKAKKLOQPSQLELYCYLYEWEEDFVORANDYFPKIEI-NLSTRMDH	661	FT	/note= "P-loop"
QY	621	MYSSFCIKRCRSAQVHLXYGATYSADGDEDRARCSAGAHTLLVQLRPETVLLDAYSEHLA	680	FT	/note= "ATP/GTP-binding site motif A (P-loop)"
Db	662	MYSSFCIENCHRVESLSL-GFLHNMPKEEBEKEGRHLDVQ-----	703	FT	/label= N-myristoylation_site
QY	681	ALCTPNLIELSLYNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSSACEDLSAAL	740	FT	/note= "Expanded nucleotide binding site domain signature motif #2"
Db	704	---CVLPS-----SSHAACSHG-----LVNSHLTSSFCRGLFSVL	735	FT	/label= Nucelotide_binding_site_domain
QY	741	IANKNLTRMDLSCNGVGFPGMMLLCBGLRHPQCRLOMIQLRKCQLESAGCAQEWASVLGTN	800	FT	/label= Protein_kinase_C_phosphorylation_site
Db	736	STQSITELDSLSNSIGDPQMRVLCVTLQHPGQNIIRLWMLGRGLSHECCFDLSVLSSN	795	FT	/label= Casein_kinase_II_phosphorylation_site
QY	801	PHLVELDLTGNALDELGLRLCOGLRHPVCRRLTWLWKICRLTAAACDELASTLVNQSL	860	FT	/note= "Expanded nucleotide binding site domain signature motif #3"
Db	796	QKLVLDLSNALGDFGIRLLCVGLXHLCLNKLMLVSCCLTSACCQDLASVLSTSHSL	855	FT	/label= Kinase_2_domain
QY	861	RELDLSINELGLGVLLCBGLRHPCTKLTBLGICRLGSAACEGLSVVLQANHLREL	920	FT	/note= "Walker B-box"
Db	856	TRLVYGENALGDSGVAILECAKPNQCNLOKGLVNSGLTSVCCSALSSVLTNQNLTHL	915	FT	/label= "CAMP/cGMP-dependent protein kinase phosphorylation site"
QY	921	DLSFNDLGDWGLWLLAAGLQHPACRLQKWLDSGLTAKACENLYFTLGINQTLTDLILT	980	FT	/note= "RGD cell attachment sequence"
Db	916	YLRNTGLDGKIKLCEGLLHPCKQLVLELDNCLTSHCCWDLSTLTSSQSRLKSLG	975	FT	/note= "Expanded nucleotide binding site domain signature motif #4"
QY	981	NNALGDTGVRELLCKRLSHPCCKLRVLWLFQMDLNKMTSHSLAALRVTKPYLDI	1033	FT	/label= Kinase_3a_domain
Db	976	NNDLGLGVNMFCEVLKQQSCLLQNLGLSEMYFNETKSALETLOBEKPELTV	1028	FT	/label= Protein_kinase_C_phosphorylation_site
RESULT 11					
ID	ABU63315				/note= "Expanded nucleotide binding site domain signature motif #5"
XX	AC	ABU63315 standard; protein; 1034 AA.			/label= Protein_kinase_C_phosphorylation_site
XX	AC	ABU63315;			/label= Protein_kinase_C_phosphorylation_site
XX	XX	23-SEP-2003 (first entry)			/note= "Dileucine motif"
DT	XX	Human pyrin domain family protein PYRIN-1.			/note= "Expanded nucleotide binding site domain signature motif #6"
DE	XX	Human; pyrin domain family; PYRIN-1; cellular differentiation; NF-kappaB;			/label= N-myristoylation_site
KW	XX	tissue typing; cellular proliferation; cell survival; apoptosis disorder;			/label= Casein_kinase_II_phosphorylation_site
KW	XX	inflammatory disorder; apoptosis associated speck like protein; ASC;			/note= "CAMP/cGMP-dependent protein kinase phosphorylation site"
KW	XX	nuclear factor kappaB.			
OS	XX	Homo sapiens.			
PH	Key	Location/Qualifiers			
FT	Domain	1..87			/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	/label= Pyrin_domain			/note= "Dileucine motif"
FT	Modified-site	3..5			/note= "Dileucine motif"
FT	Modified-site	/label= Protein_kinase_C_phosphorylation_site			/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	44..46			/note= "CAMP/cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	/label= Protein_kinase_C_phosphorylation_site			
FT	Modified-site	93..98			
FT	Modified-site	/label= N-myristoylation_site			
FT	Modified-site	110..113			

QY 561 KMDLLQWIOSKAGSDGSLTQSGSLRPFSCLYEIOEEFIQOALSHFOVIVWSNIASKMEH 620
 Db 603 RLELKWIEVKAKAKLQPSOLEFYCLYEMQEDFVRAMWYFKIEI-NLSTRMDH 661
 QY 621 MYSSFCIKRCRGAQVHLHYGATYSADGEDRARCAGAHNTLLVQLRPERTVLLDAYSEHLA 680
 Db 662 MYSSFCIENCHRVESLSL-GFLHNMPEKEEKEGRHLDWQ----- 703
 QY 681 ALCTNPNIIELSLYNALGSRGVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAAL 740
 Db 704 ---CVLPFS-----SSHAACSHG-----LVNGLHTSSFCRGLFSVL 735
 QY 741 IANKNLTRMDLSGNGVGFPMMLLCGLRHPQCRLQMIOLRKCQLESAGCOEMASVLTGN 800
 Db 736 STSQSTELDSNLSGDPQMRVLCITLQHPGNCIRRLWLGRCLGSHCECFDISLVSSN 795
 QY 801 PHVELDLTGALEDLGLRLCCGLRHPVCRRLTLWLKICRLTAACADELASTLSVNSQL 860
 Db 796 QKVELDLSDNALDGFIRLLCVGLKHLNKLMLVSCCLTSACCQDLASVLSHSL 855
 QY 861 RELDLSNLGLDGLVLLCEGLRHPCKLOTLGLICRLGSAACEGLSVLQANHNLREL 920
 Db 856 TRYVGENALGDSGVAILCEKAKNPQCNLQKGLVNSGLTSVCCSALSSVLTNQNLTHL 915
 QY 921 DLSFNDLGDWGLWLLAEGLOHPACRLQKWLDCGLTAKACENLYFTLGINQTLTDLTLYT 980
 Db 916 YLRGNTLGDGKIGKLLCEGLLHPCKLOVLELNCNLTSHCCWDLSTLLTSSQSLRKLISLG 975
 QY 981 NNALGDTGVRLCKRLSHPCCKRLVLWLFQMDLNKMTSHRLAALRVTKPYLDI 1033
 Db 976 NNDLGLGVMMFCEVLKQSQCLLQNLGLSEMYENYETKSALETLOKEKPELTV 1028

RESULT 12

ABB77910

ID ABB77910 standard; protein; 449 AA.

AC ABB77910;

XX 07-OCT-2002 (first entry)

DT Amino acid sequence of human leucine-rich repeat protein HLRBEM1.

XX Human: leucine-rich repeat; HLRBEM1; proliferative disorder;

XX immune condition; apoptosis; signal transduction; autoimmune disease;

XX haematopoietic cell disease; graft-versus-host disease; allergy; asthma;

XX cardiovascular disorder; neurological disease; pheromone;

XX pulmonary disease; chronic obstructive pulmonary disease;

XX allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;

XX haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;

XX inflammatory disorder; systemic lupus erythematosus;

XX cardiovascular disease; cancer.

XX Homo sapiens.

OS Key Location/Qualifiers

XX Domain 144..159

FT /note= "transmembrane domain"

FT

XX WO200252011-A2.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-US049740.

XX 22-DEC-2000; 2000US-0257773P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder J, Ramanathan C, Mintier G;

XX WPI; 2002-566676/60.

DR N-PSDB; ABL59333.

XX New HLRBEM1 nucleic acids for preventing, treating or ameliorating e.g.

PT proliferative disorders, immune conditions, a disorder related to

PT aberrant apoptosis modulation or developmental disorders.

XX Claim 20; Fig 1A-E; 371pp; English.

XX The present sequence represents a human leucine-rich repeat containing

CC protein, designated HLRBEM1. HLRBEM1 polypeptides and polynucleotides are

CC useful for preventing, treating or ameliorating a medical condition such

CC as a proliferative disorder, immune condition, or a disorder related to

CC aberrant apoptosis modulation, either directly or indirectly, and in

CC modulating signal transduction activity in various cells, tissue and

CC organisms. They are also useful for treating, preventing, or diagnosing

CC diseases of haematopoietic cells, autoimmune disease, graft-versus-host

CC disease, allergic conditions (e.g. asthma), cardiovascular disorders, and

CC neurological diseases, and for increasing the organisms' ability to

CC synthesize and/or release pheromones. The polypeptide may also be used in

CC treating, preventing or ameliorating pulmonary disease (e.g. chronic

CC obstructive pulmonary disease, allergic rhinitis, or bronchial

CC hyperresponsiveness), reproductive disease, haematopoietic disease,

CC platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious

CC disorders (e.g. innate immunity to bacterial pathogens, or adaptive

CC immune response), immune and inflammatory disorders, or systemic lupus

CC erythematosus), cardiovascular diseases and cancers. HLRBEM1 nucleic

CC acids may further be used in chromosome identification or mapping, as a

CC chromosome marker, as molecular weight markers, as diagnostic probes, in

CC gene therapy, in raising anti-DNA antibodies, or as antigens for

CC eliciting immune responses

XX Sequence 449 AA;

XX Query Match 42.7%; Score 2334; DB 5; Length 449;

XX Best Local Similarity 90.7%; Pred. No. 2e-214; 0; Indels 46; Gaps 2;

XX Matches 449; Conservative 0; Mismatches

QY 228 MNQATSCSMQDLIFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPHDPQGWCLCW 287

Db 1 MNQATSCSMQDLIFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPHDPQGWCLCW 60

QY 288 EKRPTELLNSLRKLLPELSLITTRPTALEKHLRHLHPHVEILGFSAEKKEYF 347

Db 61 EKRPTELLNSLRKLLPELSLITTRPTALEKHLRHLHPHVEILGFSAEKKEYF 120

QY 348 KYFHNAGAGQVFNVRDNEPLFTWCFVPLVCWVCTCIQQOLEGGGLLRQTSRTTAV 407

Db 121 KYFHNAGAGQVFNVRDNEPLFTWCFVPLVCWVCTCIQQOLEGGGLLRQTSRTTAV 180

QY 408 YMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLWNQKILPEEQDLRKHGLDGEDVSF 467

Db 181 YMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLWNQKILPEEQDLRKHGLDGEDVSF 240

QY 468 LNMNIFQKIDINCERYYSFIHLSPQEFFAAMYIYLDGEGGAGDPQDVTLLTETAYASERS 527

Db 241 LNMNIFQKIDINC-----ERS 255

QY 528 FLATSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAGSDGSLTQSGSLEFF 587

Db 256 FLATSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAGSDGSLTQSGSLEFF 315

QY 588 SCLEYIOEEFEIQQALSHPFQVIVVSNIAKMEHWSFCLKRCRQAQVHLHYGATYSADG 647

Db 316 SCLEYIOEEFEIQQALSHPFQVIVVSNIAKMEHWSFCLKRCRQAQVHLHYGATYSADG 375

QY 648 EDARCSAGAHNTLLVQLRPERTVLLDAYSEHLAAALCTNPNIIELSLYNALGSRGVKLL 707

Db 376 EDARCSAGAHNTLLVQLRPERTVLLDAYSEHLAAALCTNPNIIELSLYNALGSRGVKLL 434

QY 708 CQGLRHPNCKLQNLRL 722

Db 435 CQGLRHPNCKLQNLRL 449

RESULT 13
ADE36452
ID ADE36452 standard; protein; 719 AA.
XX
AC ADE36452;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human PAAD domain associated protein.
XX
KW cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
KW antiatheric; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; pyrin 2; PAAD domain.
XX
OS Homo sapiens.
XX
FN US2003077699-A1.
XX
PD 24-APR-2003.
XX
PF 25-SEP-2001; 2001US-00965621.
XX
PR 26-SEP-2000; 2000US-00671760.
PR 26-SEP-2000; 2000US-0367367P.
XX
PA (REED/) REED J C.
PA (GODZ/) GODZIK A.
PA (CHUZ/) CHU Z.
PA (PAWL/) PAWLOWSKI K.
PA (FIOR/) FIORENTINO L.
PA (ARIZ/) ARIZA M E.
PA (STEH/) STEHLIK C.
XX
PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
PI Stehlik C;
XX
WPI; 2002-471256/50.
DR N-PSDB; ADE36451.
XX
PT Novel isolated PAAD domain containing polypeptide useful for inducing
PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
PT therapy for treating cancer.
XX
PS Disclosure; SEQ ID NO 59; 93pp; English.
XX
CC The invention describes an isolated PAAD domain containing polypeptide
CC (I) comprising 80% identity to the amino acid sequence of PAAD and
CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
CC -like protein containing a caspase recruitment domain (ASC)-2 fully
CC defined in specification, where (I) is biologically active. (I) is useful
CC for identifying a (I)-associated polypeptide, an agent altering that
CC association and agents that modulate PAAD domain mediated inhibition of
CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
CC for identifying an agent that modulates the activity of the NB-ARC domain
CC of (I). (I) or its functional fragments is useful in altering cellular or
CC biochemical process such as apoptosis, NFkappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c

CC -Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (I) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
CC therapy. This is the amino acid sequence of human PAAD domain associated
CC protein.
XX
SQ Sequence 719 AA;
Query Match 30.2%; Score 1653.5; DB 5; Length 719;
Best Local Similarity 46.7%; Pred. No. 8.5e-149;
Matches 335; Conservative 125; Mismatches 174; Indels 83; Gaps 10;
QY 11 CELSTYLELENAVELKKKLYLGATTELGEGKIPW----- 45
DB 8 CXLARYLEDELDVLUKTFMHELYPP-QKGCIPLPXXXXXXXKXXXXXXXXXXXXXXX 66
QY 46 -----GSMKAGAPLEMAQLLITHPGPEAWRLALS 75
DB 67 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXSDNARVSNFTVICO-----EDSIEEW---MG 119
QY 76 TFERINRDXLMERGOREDLVRDPQBYTYRKYVRRKFLMEDRNARLGEVNLSHRYTRLL 135
DB 120 LLEYLSRISICK-----MKDYRKYKYVRSRQCTEDRNARLGESVSLNKRYTRLRL 173
QY 136 VKEHNPQVQOQLDGTGRHARTVGHQASPTIKITLPEPDEERPEPTVMOGAAGTG 195
DB 174 IHEHRSQOEREQELLAIGK--TKTCESVPSPITKMLLLFPDDEHSEPHVHTVFOGAAGTG 231
QY 196 KSMIAHKVMDWADGKLFQGRFDYLYINCRMNQSAECMSMODLIFSCWPPSPAPLOBL 255
DB 232 KTLARKMWDWASGTLVQDFDYLYFTHCRVS-LVTQSRSLGDLIMSCCPDPNPIHKI 290
QY 256 IRVPERLLFIIDGDELKPSFHDQPMWCLCWEKRPTELLNLSIRKLLPELSLITTT 315
DB 291 VRKPSRILFLMDGDFELQGFDEHIGPLCTDQKAERGDIILLSSLRKLLPEASLITTT 350
QY 316 RPTALEKHLRLLEHPRHVEILGFEAEKYEYKYFHNAEQAGOVNYVRDNEPLFTMCF 375
DB 351 RPALEKLOHLLDHPHVEILGFEAEKYEYKYFHNAEQAGOVNYVRDNEPLFTMCF 410
QY 376 VPLVCMVWCTCLOQLGEGGLRQTSRTTAVYMLYLLSLMQPKGAPRLQPPNQRGLC 435
DB 411 IFLVCMVWCTGLKQOMESGKSLAQTSKITTAVYVFLSLSLQPRGSGHGLCAHLWGLC 470
QY 436 SLAADGLNNQKILFBEQDLRKHGLDGVSAFLNNMIFOKDINCERYYSFIHLFOEPFA 495
DB 471 SLAADGIWNQKILFBEEDLRNHLGKADVSAPFLNNLFOKEVDCFKFYFIHMTFOEPFA 530
QY 496 AMYVILDEGEGAG-----PDQVTRLLTAYAFSERSFALTSRFLGLLNEETS 546
DB 531 AMYVILDEEGERTNVPGRSLKPLSRDVTLLNKGKFKGYLIFVVRFLGLVNGERTS 590
QY 547 HLEKSLCWKVSPIHMDLQIOWSKAQSDGSTLQOQSLEFFSCLYEIOEEFTQQALSHF 606
DB 591 YLEKLSCKISQIIRLELLKLVKAKAKKLOIQSLELFYCLYEMQEDFVQRAWDYF 650
QY 607 QVIVVSNITASKMEHWSFCLKRCRKAQVLIHUYGATYSADGEDRARCAGATHLIVQ 663
DB 651 PKIEI-NLSTRMDHWSVSCFNCIENCHRVESLSL-GFLHNPKEEBEEKEGRHLDVMQ 705
RESULT 14
AAE21062
ID AAE21062 standard; protein; 344 AA.
XX

AC AAE21062;
 XX 01-JUL-2002 (first entry)
 DT rno (upregulated by nitric oxide)-1 protein.
 DE Upregulated by nitric oxide; NO; rno-1; differentiation; death; cancer;
 XX acute myeloid leukaemia; AML; cytostatic.
 KW Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 249 /label= Unknown
 FT /note= "This residue is represented in the specification
 FT as O"
 FT Misc-difference 256 /label= Unknown
 FT /note= "This residue is represented in the specification
 FT as O"
 FT US2001029033-A1.
 XX 11-OCT-2001.
 XX 06-MAR-2001; 2001US-00799983.
 XX 06-MAR-2000; 2000US-0186971P.
 XX (SHAM/) SHAMI P J.
 XX (PARK/) PARKER C J.
 XX Shami PJ, Parker CJ;
 PI WPI; 2002-009982/01.
 XX Novel polypeptide regulated by nitric oxide useful for inducing acute
 XX myeloid leukemia cell differentiation and apoptosis.
 XX Claim 29; Fig 7B; 23pp; English.
 XX The present invention relates to novel genes which are upregulated by
 XX nitric oxide (NO), designated as rno and their corresponding proteins.
 XX The invention also relates to the isolation and characterisation of three
 XX isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
 XX are useful for inducing differentiation and death in a cancer cell.
 XX Diagnosing acute myeloid leukaemia (AML) comprises detecting the
 XX expression of a rno gene comprising the nucleotides coding for rno
 XX protein. The present sequence is rno-1 protein
 XX Sequence 344 AA;
 SQ
 Query Match 30.2%; Score 1652; DB 5; Length 344;
 Best Local Similarity 96.9%; Pred. No. 3.6e-149;
 Matches 316; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
 QY 710 GLRHPCNQLRKRKCRSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 769
 DB 21 GLQOP--RRRLWKRCRISACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 78
 QY 770 HPQCRILQMIQLRKQLESGACQEMASVILGNPHLVLDLTGNALDGLRLCCQLRHVP 829
 DB 79 HPQCRILQMIQLRKQLESGACQEMASVILGNPHLVLDLTGNALDGLRLCCQLRHVP 138
 QY 830 CRURTWLKICRTAAACDELASTLSVNSQIRRELDLSINELGDLGVLLCEGLRHPTCKL 889
 DB 139 CRURTWLKICRTAAACDELASTLSVNSQIRRELDLSINELGDLGVLLCEGLRHPTCKL 198
 QY 890 QTLRLGICRGSACCEGLSVVLQANHLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKL 949
 DB 199 QTLRLGICRGSACCEGLSVVLQANHLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKL 258
 QY 950 WLDSCGLTAKACENLYFTLGINTLDTLNTNMGDTGVRLLCKRLSHPGCKLRVLWLF 1009

DB 259 WLDSCGLTAKACENLYFTLGINTLDTLNTNMGDTGVRLLCKRLSHPGCKLRVLWLF 318
 QY 1010 GMDLNKMTSHRLAALRVTKPYLDIGC 1035
 DB 319 GMDLNKMTSHRLAALRVTKPYLDIGC 344
 RESULT 15
 AAO15593
 ID AAO15593 standard; protein; 994 AA.
 XX
 AC AAO15593;
 XX 31-OCT-2002 (first entry)
 DT Human PYRIN-3 protein.
 DE Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-3.
 XX Homo sapiens.
 OS
 XX WO200261049-A2.
 XX 08-AUG-2002.
 XX 31-JAN-2002; 2002WO-US002967.
 XX 31-JAN-2001; 2001US-0265231P.
 XX 10-SEP-2001; 2001US-0318645P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX (AMHP) WYETH.
 XX Bertin J, Wang W, Blatcher M;
 PI WPI; 2002-627477/67.
 XX N-PSDB; AAL44366.
 XX New PYRIN polypeptides and nucleic acids useful for modulating and
 XX diagnosing stress-related, apoptotic and inflammatory responses, or for
 XX treating inflammatory and immune system disorders, cancers, or
 XX neurological diseases.
 XX Claim 8; Fig 11; 167pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 XX useful for modulating and diagnosing stress-related, apoptotic and
 XX inflammatory responses. The PYRIN protein and DNA sequences are useful
 XX for treating inflammatory disorders and immune system disorders (e.g.
 XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
 XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
 XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
 XX (e.g. systemic lupus erythematosus and arthritis); and neurological
 XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 XX protein and DNA sequences may also be used in screening assays, detection
 XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 XX predictive medicine (e.g. diagnostic assays, clinical trials and
 XX pharmacogenomics) and transcription profiling. The present amino acid
 XX sequence represents the human PYRIN-3 protein
 XX Sequence 994 AA;
 SQ
 Query Match 26.6%; Score 1457.5; DB 5; Length 994;
 Best Local Similarity 33.9%; Pred. No. 9.2e-130;
 Matches 353; Conservative 197; Mismatches 402; Indels 89; Gaps 19;

Search completed: July 30, 2004, 13:49:22
Job time : 61 secs

QY 13 LSTYLELEAVELEKFKLYLGAT- ELGEGKI PKGMEKAGLEMAOQLLITHFGPERAWR 71
Db 11 LMWYLEELKEEFKFKHEHLKQMTLQELKQIPWTEVKASREELANLLIKHYEEQAWN 70
QY 72 LALSTFRIRNKDILWEGQREDLVRDPOETRYDYVRKPF-RLMEDRNARLGEVNLSHRY 130
Db 71 ITLAIFQKMDKDLCKVMRER--TGYTKTYAHAKQKFSLSWSSKSV----- 116
QY 131 TRLLLVKHSNPMOVQOQLDTRGHARTVGHQASPIKIETLFPDDEERPEPPRTVMQ 190
Db 117 TEIHLVFEE---EVKOECD-----HLDRLFAP-KETGKQPRTVIIQG 155
QY 191 AAGIGKSLAHKVMLDWADGKLFQGRDYLYINCREMNQSAECSDMODLIFSCKWPEPSA 250
Db 156 POGIGKTTLLMKLWMSNDKIFRDFLYTYFCCREURE-LPPTSADLISREWPOFAA 214
QY 251 PLOBLIRVPERLLFTIDFDELKPSFHDPPQGPWCLCWEKRPTELLNLSLRKLLPELS 310
Db 215 PITEIVSQPERLLFVIDSFEELQGLNEPDSDLCDLMEKRPVQVLLSSLLRKKMLPEAS 274
QY 311 LLITRPTALEKHLRHLHLEPHRVEILGFSERKEKEYFYKYPHNAEQAGQVNYVRDNEPL 370
Db 275 LLIAIKVPCKELRDQVTISIIYQPRGFNEDRLVYFCCFPKPKRAEAFNLVRESQ 334
QY 371 PTMCFVPLVCHWCTCLOOQLEGGGLRQTSRTTAVNMLYLLSLMQPKGA--PRLOPP 428
Db 335 FSIQOILPCLWILCTSKQEMQKDKALTCOSTTSVSSVFNLTFPE-GAEGFTPTQ 393
QY 429 PNQRLGSLAADGLWNOKILFEEQDLKXHGDLGDEVSFAFLNMNIFQKINCERYYSFIHL 488
Db 394 HOLKALCSLAEGNWTDTFECEDDLRENGVVDADIPALLGTILLKYGERESSVFLHV 453
QY 489 SPQEFFAAMYI-----LDEGGGAGPDQDVTRLTEYAFSERSFALTSRFLGLLNEET 544
Db 454 CIOEFCAALFYLLKSHLDHPHAPVRCVOEL--LVANFEKARRAHWIFLGCELTGLLNKKE 511
QY 545 RSHLEKSLCWMKVSPIHMDLLOWIOSKAQSDGSTLQOGLSEFFSCLYEIOEEETQOALS 604
Db 512 QEKLDAPFGLOLSEIQQIHHQCLKSGERENPQGVDSLAIFYCLFEMQDPAPFKQAVN 571
QY 605 HFQVI--VVSNIASKMEHMYSSFCLEKRCRAQVILHYGATYSADGEDRARCASAGHTIL 661
Db 572 LLOEANEHIIIDNV---DLVVSAYCLKYCSSURKLCF--SVQNVFKKEDHSSTSDYSLI 625
QY 662 VOLRPERTVLDAYSEHIAALCTNPNIELSLYRNALSGRGVLLCOGLRHPNCKLQNL 721
Db 626 CW-----HHICSVLTTSGHLRELQVQDSTLSESTFVTWCNQLRHPSCRLQKL 672
QY 722 RLKRCRISSAC-----EDLSAALIANKLTRMDLSGNGVGFPGMMLLCGLRHPQ 772
Db 673 GINNVSFGQSVLLFEVLYQPDULKYLSFTTKLSRDDIRS-----LCDALNYP 722
QY 773 CRLQMIQLRKQOLESAGCEMASVLGNPHILVELDTGNALEDLGLRLCCGLRHPVCR 832
Db 723 GNVKELALVNCHLSPIDCEVLGLLITNNKLTYNVSCNQL-DTGVPLLCALCSPDVL 781
QY 833 RTLWIKICRLTAACDEBLASTLSVNQSLRELDLSNELGDLGVLLLCGLRHPCKLOT 892
Db 782 VYLMALFCHLSEQCEYISEMLLRNKSRYLDLSANVLKDEGLKTLCEALKHPDCCDLS 841
QY 893 RLIGICRLGSAACEGLSVVLOANHNRLDLSFNDLDGMDGLMLAEGLOHPACRLQKMLD 952
Db 842 CLVKCFITPAACEDLASALISQNLKILOIGCNEIGDVGVQLLCALTHTCRLEIGLE 901
QY 953 SCGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLCKRISHPGCKLRVLWLFMD 1012
Db 902 ECGLTSTCKDLASVLTCSTKIQLNLTLNTLTDHTGVVVLCEALRHPICALQVLGLRKT 961
QY 1013 LNKWTHSRILAARVTKPYLDI 1033
Db 962 FDEETQALLTAEERPNPLTI 982

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:45:21 ; Search time 22 Seconds

(without alignments)

4525.374 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472

Sequence: 1 MLRTAGRGCLSLTYLEEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314.5	24.0	1192	2 T17255	hypothetical prote
2	1160.5	21.2	1111	2 A59000	mater protein [imp
3	724	13.2	461	2 A31858	ribonuclease-angio
4	717	13.1	456	2 S20597	ribonuclease inhib
5	709	13.0	456	2 A31857	ribonuclease inhib
6	494.5	9.0	483	2 S27880	Nasopressin recept
7	377.5	6.9	1130	2 A48843	MHC class II trans
8	258	4.7	1004	2 T31665	hypothetical prote
9	216.5	4.0	1075	2 T31668	hypothetical prote
10	183.5	3.4	1232	2 A55478	neuronal apoptosis
11	180	3.3	312	2 B97746	hypothetical prote
12	179.5	3.3	545	2 T52068	RAN GTPase-activat
13	178	3.3	533	2 T52063	ran GTPase-activat
14	177.5	3.2	506	2 A45841	T-complex-associat
15	172.5	3.2	618	2 T48193	hypothetical prote
16	171	3.1	1447	2 T42628	neuronal apoptosis
17	170	3.1	568	2 F86291	hypothetical prote
18	169.5	3.1	789	2 T52067	hypothetical prote
19	169.5	3.1	1121	2 T02764	myosin-I binding p
20	162	3.0	631	2 C89243	protein F28C1.3 [i
21	162	3.0	631	2 T21471	hypothetical prote
22	161.5	3.0	589	2 A36983	RNAI homolog fugi
23	160	2.9	998	2 T23427	hypothetical prote
24	158	2.9	535	2 T48102	RAN GTPase activat
25	156.5	2.9	589	2 T52070	RNAI protein homol
26	155	2.8	587	2 JC5300	Ran GTPase activat
27	152.5	2.8	2493	2 A55481	adenylate cyclase
28	152	2.8	1010	2 T36383	probable large ATP
29	142.5	2.6	1253	2 T45787	disease resistance

30	142	2.6	934	2 T05201	hypothetical prote
31	140.5	2.6	1389	2 T13852	gene wheeler prote
32	140	2.6	1039	2 A85096	hypothetical prote
33	139.5	2.5	271	2 D84586	hypothetical prote
34	138	2.5	526	2 C84552	hypothetical prote
35	137.5	2.5	1237	2 A81915	hypothetical prote
36	137.5	2.5	1385	2 T13887	tlr protein - frui
37	136.5	2.5	1286	2 T33476	hypothetical prote
38	135.5	2.5	1066	2 T15864	hypothetical prote
39	135	2.5	592	2 T52139	LRR-containing F-b
40	135	2.5	1217	2 T52348	disease resistance
41	134.5	2.5	906	2 G96621	probable disease r
42	134.5	2.5	957	2 E84547	probable disease r
43	134	2.4	607	2 E96598	protein F20N2.2 [i
44	133	2.4	1110	2 F84547	probable disease r
45	132.5	2.4	907	2 J80176	orphan G protein-c

ALIGNMENTS

RESULT 1

T17255

hypothetical protein DKFp586O1822.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17255

R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18722

A:Accession: T17255

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1192 <ROE>

A:Cross-references: EMBL:AL117470

A:Experimental source: adult uterus; clone DKFp586O1822

C:Genetics:

A:Note: DKFp586O1822.1

Query Match 24.0%; Score 1314.5; DB 2; Length 1192;
Best Local Similarity 39.9%; Pred. No. 4.6e-94;
Matches 323; Conservative 119; Mismatches 293; Indels 75; Gaps 19;

QY	133	LLLVKEH---SNPMQVQO--QLLDTGRGHARTVGHQASPIKIEITLFEDEERPEPRITVV	187
DB	1	LLLQRPSPRQDPLVKRSWPDYVEENRGL-----IEIRDLFGPLDQEQ-PRIVI	50
QY	188	MQGAAGICKSMLAHKVMLDWADGKLFQGRFDYLFYINCRMNQOSATECSMODLIFSCWPE	247
DB	51	LQGAAGICKSTLARQVKEAWRGQLYGDRFQHVFFSCRELAQSKV-VSLAELIGKDGTA	109
QY	248	PSAPIQLIRVPERLLFIIDGFDLKPSPHDPQGPWCWCWEKRPTEILLNSLRKGLP	307
DB	110	TPAPIRQILSRPERLLFIIDGVDPEPGVQLQEPSSSELCLHWSQPQADALLGSLGKLTLP	169
QY	308	ELSLITTPALEKHLRLLHPRVEILGSEAEKKEYFYKFNACQGVFNVRDN	367
DB	170	EASFUIITARTALQNLIPSEQARWVEVLGFSSSRKKEYFYFYFTDQRAIRAFRLVSN	229
QY	368	EPLFTMCVPVLVWVCTCLOQLGEGGLLRTSTTTAVVLYLLSLMQPKGAPRLQP	427
DB	230	KELWALCLVFWVSWLACTCLMQMKRKEKLITLSTTTTCLHYLAQAQLQPLGQL---	287
QY	428	PPNQGLCSLAADGLWNOKILFEEDLRKHGDLGEDVSFAFLNMNIFQKDKNCERYYSFIH	487
DB	288	---RDLCSLAAGIWKTKTLFSPDDLKRKHGDLGAIISTFLKMGILQEH-PIPLSYSFIH	342
QY	488	LSFQEFFRANVYILDEGGGAGPDQ----DVTRLLTYAFPSERSPLATTSRFLGLNNEE	543
DB	343	LCFQEFFAAMSIVL--EDEKGRGKHSNCITIDLEKTLAEYGI-HGLFGASTTRFLGLLSDE	400
QY	544	TRSHLEKSLCWKSPHKNMDLLQWIOSKAQSDGSLTQQGSLEFFSCFLYEIOEEETIQAL	603

Db 401 GEREMENIFHCRLSQ--GRNLQWVPSLQ-----LLQPHSLBESLHCLYETRNKTELQVM 454
 QY 604 SHFQVIVSNIAKMEHNVSSCLKRCRQAQVHLGYATYSADGEDRARCAGATILVQ 663
 Db 455 AHFEEMGMC-VETDMELLVCTCFIKFSRVKQLI-----BGRQHRSTWSPMTMVL 505
 QY 664 LPERTVLLDAYSEHLAAALCTNPNIIELSLYRNALSGVGLKLCQGLRHPNCKLQNLRL 723
 Db 506 FR-WPVVTDATQWILFVLKTRNLKELDLNSLSHSAVSLCKTLRRPCLLETLEL 563
 QY 724 KRCRISSACEDLSAALANKNLTRMDLSNGVGFPGMMLLCEGLRHPQCRLOMIQLKRC 783
 Db 564 AGCGLTAEDCKDLAFGLRANQTLTDLDFNVLTDAKAKHLQRLQPSCKLQRLQVSC 623
 QY 784 QUESACQEMASVGLTNPVPHVLDLTGNALEDGLRLCQGLRHPVCRILRTWLKICRLT 843
 Db 624 GLTSDCCODLASVLSASPSSLKELDLQNNLDDVGVLELCEGLRHPACKLIRLGL----- 677
 QY 844 AAACDELASTLSVNOQLRELDLSNEL-----GDLGVLLLCGL-----RHPTCKLQTLR 893
 Db 678 ----DQTLSDMRQELRALQEKPOLLIFFSRKPSVMTPTTEGLDITGEMNSSTSLKQR 733
 QY 894 LGICRLGSAACEGLSVVLOANHLRELDLS 923
 Db 734 LGSERAASHV-----AQANLKLDDVS 754

 RESULT 2
 A59000
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: A59000
 R:Tong, Z.B.; Nelson, L.M.
 Endocrinology 140, 3720-3726, 1999
 A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ovulation
 A:Reference number: A59000; MUID:99360614; PMID:10433232
 A:Accession: A59000
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1111 <KUR>
 A:Cross-references: GB:AF074018; NID:g5802697; PIDN:AAD51762.1; PID:g5802698
 A:Genetics:
 A:Gene: Mater

 Query Match 21.2%; Score 1160.5; DB 2; Length 1111;
 Best Local Similarity 31.4%; Pred. No. 4.6e-82;
 Matches 293; Conservative 165; Mismatches 407; Indels 69; Gaps 13;

 QY 162 HQASP-1K-IETLFEPEDEPEPPRTVVMQGAAGIGKSLAHKVMLDWADGKLFQGRFDY 219
 Db 168 HYDSPENKLLSDAFKP-YQKTFQPHIILHGRFGVGSALARSIVLQWAGKLFQ-KMSF 225

 QY 220 LFYINCREMNQSAECMODLIFSCWPEPSAPIQELIRVPERLLFIIDGDELKPSFHPD 279
 Db 226 VIFSVREIKWT-EKSLAQIAKECPDSMDLVTKMSQPERLLFVLDGLDDMSVLQHD 284

 QY 280 QGPWCLCWEKRTPELLLSLIPKLLPELSLITTPRTALEKHLRHPHVEILGFS 339
 Db 285 DMTLSRDWQDEQPIYILMYSLLRALLPQSLFIITRNTGLEKLSKMSVSPVLVEGLS 344

 QY 340 EAEKKEVYFYKPHNAEQAGVFNVRDNEPLFTWCFVPLVWVCTCLOOQLEGGLLRQ 399
 Db 345 ASRSQVLNENISNESDRIQVHSLIENHQLFQCCQAFSVCSLVCEALQKGLKRCRTL 404

 QY 400 TSETTAVY---MVLISLMQPKGAPRIQPPNQRLGSLADGLMNQKILFEEDQLRK 456
 Db 405 PCQTLGLVATLVFHLQTLKRPQSALSQBEQITLVGLCMAAEGVMTWFSVFYDDDKN 464

 QY 457 HGLDGDVSAFANMT-FOKDINCERYVSIPIHLSFOFFAAMYIILDEGGGAGPDQ--- 512
 Db 465 YSLKESEILAFHNNILQVGHNSQCYVFSLSLQDFFAALYYVL---EGLEWNOHFC 521

QY 513 --DVTRLLTEYAFSERSFALTSRFLGLLNEETSHLEKSLCWKVSPIHKMDLQWIOS 570
 Db 522 FIENQRSIMEVKRTDTRLLGKMKRFLGLMKNKDLIKTLEVLPEYVPIPTVEQLQHWVSL 581

 QY 571 KAQSGSTLQOQSLRFFSCLYIBIEEFIQOALSHFOVIVSNIAKMEHNVSSCLKRC 630
 Db 582 IAQVNGTSPMDTLDAFYCLFESODEEFVGGALKRPFQVWLL-INQKMDLVSSYCLKHC 640

 QY 631 RSAQVHLH-----YGATYSADGE----- 648

 Db 641 QNLKAIKRVDIRLLSVNTLELCVVTVTOCKPMLMEWNGFCSVLGSLNKLKELDGL 700

 QY 649 DRARCAGAHITLIVQLRPE-----RTVLLDAYSEHLAAALCTNPNIIELSLYRNAL 699

 Db 701 DSILSORAMKILCLELRNQCRIQKLTFKSAEVSGLXHLKLLFNSQNLKYLNLGNTPM 760

 QY 700 GSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALANKNLTRMDLSNGVGF 759
 Db 761 KDDDMKLACEALKHPKCSVETLDSCELTIIYEMISTLLI-STTRKLCLSAKNRVGVK 820

 QY 760 GMLLCEGLRHPQCRLOMIQLRKKCOLESGACQEMASVLTGTPHVLVDLTGNALEDGLR 819

 Db 821 SMISGNALSSSMCLLQKILNDNCGLTPASCHLLVSALFNSQNLTHLCLSNNSLGTGVQ 880

 QY 820 LLCQGLRHPVCRILRTLWLKICRLTAAACDELASTLSVNOQLRELDLSLNLGDLGVLLC 879
 Db 881 QLCQFLRNPECALQRLILNHCNIVDDAYGFLAMRLANNTKLTHTLSLTMPVSGAMKLLC 940

 QY 880 EGLRHPTCKLQTLRLGICRLGSAACEGLSVVLOANHLRELDLSFNDLGDGWLMLAEG 939

 Db 941 EALKEPTCYLOELVDCQLTQNCCEDLACMTITTKHLKSLDLGNALGDGVITLCEGL 1000

 QY 940 QHPACRLOKLWLDSCGLTAKACENLYFTLGINQTLTDLVLTNNALGDTGVRLLCRSLHP 999

 Db 1001 KQSSSLRLRLGKACKLTNCEALSLAISCNPHLSNLVKNDFSTSGMLKLSAFQCP 1060

 QY 1000 GKCLRVLMFGMDLNKQMTSHRLAALRVTPYLDI 1033
 Db 1061 VSNLGIIGLWKQBYARVRRQLEEFVFKPHVVI 1094

 RESULT 3
 A31858
 ribonuclease-angiogenin inhibitor - human
 N:Alternate names: ribonuclease inhibitor, placental
 C:Species: Homo sapiens (man)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000
 C:Accession: A31858; S02012; S23933; S48636; T47188
 R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
 Biochemistry 27, 8545-8553, 1988
 A:Title: Primary structure of human placental ribonuclease inhibitor.
 A:Reference number: A31858; MUID:89118269; PMID:3219362
 A:Accession: A31858
 A:Molecule type: mRNA
 A:Residues: 1-461 <LEE>
 A:Cross-references: GB:M22414; NID:g186260; PIDN:AAA59130.1; PID:g307040
 R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
 EMBO J. 7, 4151-4156, 1988
 A:Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) disclosed
 A:Reference number: S02012; MUID:89210799; PMID:3243277
 A:Accession: S02012
 A:Molecule type: mRNA
 A:Residues: 1-422, 'SE', 425-461 <SCH>
 A:Cross-references: EMBL:X13973; NID:g35843; PIDN:CAA32151.1; PID:g35844
 A>Note: part of this sequence, including the carboxyl end of the mature protein, was cor
 R;Crevel-thieffry, I.; Cottierill, S.; Schuller, E.
 Biochim. Biophys. Acta 1122, 107-112, 1992
 A:Title: Characterization of a tryptic peptide from human placental ribonuclease inhibi
 A:Reference number: S23933; MUID:92338217; PMID:1633192
 A:Accession: S23933
 A:Molecule type: protein
 A:Residues: 174-195 <CRE>
 R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.

Arch. Biochem. Biophys. 312, 421-428, 1994
A;Title: Purification and characterization of human brain ribonuclease inhibitor.
A;Reference number: S48636; MUID:94311593; PMID:8037455

A;Accession: S48636
A;Status: preliminary
A;Molecule type: protein
A;Residues: 2-14 <NAD>
R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: 224374
A;Accession: T47188
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-461 <AAA>
A;Cross-references: EMBL:AL161967
A;Experimental source: adult testis; clone DKFZp434K249
C;Genetics:

A;Gene: GDB:RNH
A;Cross-references: GDB:125274; OMTM:173320
A;Map position: 11p15.5-11p15.5
A;Note: DKFZp434K249.1
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: blocked amino end; duplication
F;34-433/Region: leucine-rich 57-residue repeats
F;57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;174-195/Region: inhibitory
F;228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Query Match 13.2%; Score 724; DB 2; Length 461;
Best Local Similarity 42.5%; Pred. No. 1.7e-48;
Matches 150; Conservative 64; Mismatches 139; Indels 0; Gaps 0;

Qy 671 LLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISS 730
Db 40 LTEARCKDIIRSAIQANPALTELSLRTNELGDAGVGLVQLQGLNPTCKIQLSLQNCSLTE 99
Qy 731 SACEDELSAALIANKNLIRMDLSGNGVGPFGMMILCEGLRHPQCRLOMIQLRKQLESAC 790
Db 100 AGCGVLSLTPLTQLDELHNDNPLGDEGLKLCCEGLRDQCRLKLEKLEYSLSAASC 159
Qy 791 QEMASVLTGNTPHLVELDTGNALDGLVLLICEGLRHPVCRRLTLWLKICRLTAACDEL 850
Db 160 EPLASVLRKPDFKELVLSNNDPHEAGIHTLCQGLKDSACQLESLEKLENCGITSANCKDL 219
Qy 851 ASTLSVNQSLRELDLSNELDGLVLLICEGLRHPQCRLOMIQLRKQLESAC 910
Db 220 CGIVASKASRLRELALSGNKGVDGMAELCPGLLHFSRLTLWIWECGITAKGCGDLCEV 279
Qy 911 LQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKILWLDSCGLTAKACENLYFTLGI 970
Db 280 LRAKESLKELSLGNELDEGARLLCETLEPGCQLESLSWVKSFTAACCSHFSSVLAQ 339
Qy 971 NQTLTDLYTNALGDTGVRLLCRLSHPGCKLRVLMFGMDLNKMTWSRLAAL 1023
Db 340 NFFLELQTSNNRLDAGVRELCOGLGPGSVLRVLMWLDGCDVDSGSSCSLAA 392

RESULT 4

S20597
Ribonuclease inhibitor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C;Accession: S20597
R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992
A;Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
A;Reference number: S20597; MUID:92162755; PMID:1536887
A;Accession: S20597
A;Molecule type: mRNA
A;Residues: 1-456 <AW>

A;Cross-references: EMBL:X62528; NID:957670; PIDN:CAA44388.1; PID:957671
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.1%; Score 717; DB 2; Length 456;
Best Local Similarity 41.5%; Pred. No. 5.9e-48;
Matches 147; Conservative 66; Mismatches 141; Indels 0; Gaps 0;

Qy 671 LLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTEVRCKDIIRSAIQANPALTELSLRTNELGDAGVGLVQLQGLNPTCKIQLSLQNCSLTE 94
Qy 731 SACEDELSAALIANKNLIRMDLSGNGVGPFGMMILCEGLRHPQCRLOMIQLRKQLESAC 790
Db 95 AGCGVLPDVLRSLSLTRELHNDNPLGDEGLKLCCEGLRDQCRLKLEKLEYSLSAASC 154
Qy 791 QEMASVLTGNTPHLVELDTGNALDGLVLLICEGLRHPVCRRLTLWLKICRLTAACDEL 850
Db 155 EPLASVLRKPDFKELVLSNNDPHEAGIHTLCQGLKDSACQLESLEKLENCGITSANCKDL 214
Qy 851 ASTLSVNQSLRELDLSNELDGLVLLICEGLRHPQCRLOMIQLRKQLESAC 910
Db 215 CDVVASKASLQELDLSGNKLGNTGIAALCSGLLPSCLRLTLWLDGCDVTAEGCKDLCEV 274
Qy 911 LQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKILWLDSCGLTAKACENLYFTLGI 970
Db 275 LRAKESLKELSLGNELDEGARLLCETLEPGCQLESLSWVKSFTAACCSHFSSVLT 334
Qy 971 NQTLTDLYTNALGDTGVRLLCRLSHPGCKLRVLMFGMDLNKMTWSRLAAL 1024
Db 335 NSSLFELQMSNPLGDSGVVELCKALGYDPTVLRVLMWLDGCDVDSGSSSLATV 388

RESULT 5

A31857
Ribonuclease inhibitor, hepatic - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
C;Accession: A31857; A35830
R;Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
Biochemistry 27, 8537-8544, 1988
A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals the
A;Reference number: A31857; MUID:89118268; PMID:3219361

A;Accession: A31857
A;Molecule type: protein
A;Residues: 1-456 <HOF>
R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
Biochemistry 29, 8827-8834, 1990
A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclease
A;Reference number: A35830; MUID:91104783; PMID:2271559
A;Accession: A35830
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 82-456 <VIC>
A;Cross-references: GB:M8700; NID:9164638; PIDN:AAA63448.1; PID:9164639; GB:J02925
C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: liver
F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.0%; Score 709; DB 2; Length 456;
Best Local Similarity 42.7%; Pred. No. 2.5e-47;
Matches 151; Conservative 60; Mismatches 143; Indels 0; Gaps 0;

Qy 671 LLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTEEHCKDIGSALRANPSLTLELRTNELGDAGVHLVQLQGLSPCTCKIQLSLQNCSLTE 94
Qy 731 SACEDELSAALIANKNLIRMDLSGNGVGPFGMMILCEGLRHPQCRLOMIQLRKQLESAC 790
Db 95 AGCGVLPDVLRSLSLTRELHNDNPLGDEGLKLCCEGLRDQCRLKLEKLEYSLSAASC 154
Qy 791 QEMASVLTGNTPHLVELDTGNALDGLVLLICEGLRHPVCRRLTLWLKICRLTAACDEL 850

Db 155 EPLASVLRATRALKELTUNNDIGEGARVLGQGLADSAQOETLRLNGLTPANCKDL 214
 QY 851 ASTLSVNSQISRELDLSNELGDLVLLICEGLRHPCKLOTLRGLICRGLSAAACEGLSV 910
 Db 215 CGIVASQASRELDLSNGLGSDAGIAELCPGLSPASRLKTLWLWECDDITASGCDLRCV 274
 QY 911 LOAHNHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKLDSCGLTAKACENLYFTLGI 970
 Db 275 LOAKETLKSLAGNKLGBEGARLLCESLLOQCQLESWVKSCTTAACQHVSLMLTQ 334
 QY 971 NOTLDTLYTNALGDTGVRLLCRLKSHPCCKLRVLWLFMDLNKMTHTSRLAAL 1024
 Db 335 NKHLELQLSSNKLGDGSIQELQALSQPOTTLRLVLCGLCEVTNSGCSLASL 388

RESULT 6
 S27880
 MHC Class II transactivator - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
 C:Accession: S27880
 R:Herrera, V.L.; Ruiz-Opazo, N.
 submitted to the EMBL Data Library, February 1992
 A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
 A:Reference number: S27880
 A:Accession: S27880
 A:Molecule type: mRNA
 A:Residues: 1-483 <HER>
 A:Cross-references: EMBL:M85183; NID:g202805; PIDN:AAA03623.1; PID:g202806
 C:Genetics:
 A:Gene: AVP

Query Match 9.0%; Score 494.5; DB 2; Length 483;
 Best Local Similarity 28.8%; Pred. No. 1.6e-30; Indels 85; Gaps 10;
 Matches 147; Conservative 74; Mismatches 205;

QY 391 LEGGILLRTSTTTAVMYLLSLMQPK-PCAPRLQPPNPQGLCSLAADGLWNGKILF 449
 Db 1 MELGRDLSRTSTTISVLLFITSLKSAAGTPRVQ--GELRMLCRLAREGLKHQAQF 58
 QY 450 EQODLR--KHGLDGBDV-SAFINMIFQKINCERYSPHLSPOFFFAAMYIIDELEG 506
 Db 59 SEKDLERLKLGQSQVQTMFLSKKELPGVLETVVVYQIDQSFORFLAALSYLLD-ABG 117
 QY 507 GAGPDQVTRLLTEYAFSERSFLATSLRFLGLNEETSHLEKSLCWKVSPIKMDLLO 566
 Db 118 AFENSAGSVQMLNDSAGLRGHALTTRFLFGLLSTERTRDIGNHFGCVVPGRVQDTLR 177
 QY 567 WTQSKAQSDGSTL-----QQGSLEFFSCLYRIQESEFIQO 601
 Db 178 WYQGSQPKVATVGAEEKKDELKDEABEEEEEEELNFGLELLYCLYETQEDDFVQ 237
 QY 602 ALSHFQVIVVNI-ASKMEHWYSFCLKCRSAQVILHYGATYSADGEDRAR----- 652
 Db 238 ALSSLPFWLVRVTRMDLEVLSCYVQCPCQALRLVSCGLVAAKEKKKKKGFMMRL 297
 QY 653 ---CSAGHTLLVQLRPVTRVLLDAYSEHLAAALCTNPNIELSLRYNALSGRGVKLLCQ 709
 Db 298 KGSQSTQKPPASILLR-----LCE 317
 QY 710 GLRHNCILQNLRLKCRISACEDLSAALIANKNLRTMDLSGNGVGFPGMMLCEGLR 769
 Db 318 AMITQQCGLSILTLSHCKLPDCAVCRDLSEALKVAPSLRELQLNRLTEAGURLLSQGL 377
 QY 770 HPQCRLOMLQKCKCOLESACQEMASVLGTNPVHVELDLTGNALEDLGLRLLCQGLRHP 829
 Db 378 WPKCKVQTLRIQMPGLQE-VIHYLVIVLQSSVLTITLDSGCLPFTVVEPLCSALKHPK 436
 QY 830 CLRRTLWLKICRLTAAACDELASTLSVNQSL 860
 Db 437 CGLKTLSLTSVELTENPLRELQAVKTLKPD 467

RESULT 7
 A48843
 MHC Class II transactivator - human
 N:Alternate names: CIITA
 C:Species: Homo sapiens (man)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999
 C:Accession: A48843
 R:Steimie, V.; Otten, L.A.; Zufferey, M.; Mach, B.
 Cell 75, 135-146, 1993
 A:Title: Complementatation cloning of an MHC class II transactivator mutated in hereditary
 A:Reference number: A48843; MUID:94006536; PMID:8402893
 A:Accession: A48843
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1130 <STE>
 A:Cross-references: GB:X74301; NID:9414112; PIDN:CAAS2354.1; PID:9414113

Query Match 6.9%; Score 377.5; DB 2; Length 1130;
 Best Local Similarity 23.6%; Pred. No. 8e-21;
 Matches 220; Conservative 122; Mismatches 355; Indels 235; Gaps 40;

QY 92 EDLVDPQETY-----RDYVRKFLMEDRNARLGEVCNLSHRYTRLLLVKHSNPMQV 145
 Db 339 EQFYRLQDITYGAEPAGPDGILVEVDLVQ---ARL-----ERSSKSLERELATPDWA 388
 QY 146 CQQLDTCRGHARTVGHQASPIKIEITLPEDEERPEPPRTVVMQGAAGIGKSLAHKWL 205
 Db 389 ERQLAQGGIAEV-----LLAAKEHRRPRETRVIAVLGKAGGKSTWAGAVSR 435
 QY 206 DWADGKLFQGRFDYLYFINCREMNSQATSCSMDLIFSCWPPSPAPLOE---LIRVPER 261
 Db 436 AMACGRLPQ--YDFVFSVPCNLRPGDAYGLQDLFLSLGQPLVAADVFVSHILKRPDR 493
 QY 262 LFIIDGDELKPS---PHDQGPWCLCWEERKPTL---LNSLRKLLPELSLITTR 316
 Db 494 VLLIIDAEELEAQDGFHSTCGP-----APAEPCSLRGLLAGLFQKLLRGCTLLTAR 548
 QY 317 P-----TALEKLRLLEHPRHVEILGFSAEERKEYFYKFHNA---EQAGQVFNVDNE 368
 Db 549 PGRILVQSLSKADALF-----ELSGFSMEQQAQVVMRYFSSGTEHODRALTLLRDP 602
 QY 369 PLFTMCFVPLVWVVTCLQQOEGGGLRQTSRTTAVMYLLSLMQPKFAPRLQPP 428
 Db 603 LLLSHSPTLCRAVCQLSEALLEGEDAKLPS-TLTGLYVGLL-----GRAALDSP 653
 QY 429 PNORG-LCSLAAD-GLWQKILPEQ-----DLR-----KHGLDGEDVSAPLANNIFOKDIN 478
 Db 654 PGALAEALAKAWELGERHQSTLQEDQPPSADVYRTWAMAKGLVQHPRA----- 701
 QY 479 CERYYSFTHLSFQEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSER---SFLALTSRF 535
 Db 702 AESELAFFSFLQCFGLALWAL-----SGEIKDKELPQVYALTPRKRPYDNWLEGVPRF 757
 QY 536 LFLGLNEETRSHLEKSLCWKVSPIKMDLLOWIQS---KAQSDGSTLQQGSLEFFSCLYE 592
 Db 758 LAGLIFQPPARCLGALL--GPSAAASVDRKQKVLARYLKRQLQPTLRARQLLELHCAHE 815
 QY 593 IQEESFIIQALSHFQVIVVSNIASQMEHWVSFCLKRCESAQVILHYGATYSADGEDRAR 652
 Db 816 ABEAGIWHQ-----VWQELPGR-----SFLGTRLTTPDA-HVLGKALEAAGQD--- 858
 QY 653 CSAGHTLLVQLRPVTRVLLDAYSEHLAAALCTN--PNLIELSL---YRNALSGRGVKLL 707
 Db 859 -----FSLDLR-----STGICPSGLSGILVGLSVCTVFRFAALSD--TVAL 895
 QY 708 COGLR-HPNCKL-----QNL-RLKRCRISSECEDESALAIANK 744
 Db 896 WESLRHQETKLLQAAEBKFTIEPPKAKSKDVEDLGLVQTRTRSSSEDTAGELPAVR 955
 QY 745 NLRMDLSGNGVGFPGMMLCEGLRHPQCRQLQMIQLRKQLESACQEMASVLGTNPVHLV 804
 Db 956 DLKLEFALGPVSGQAF-----PKLV 977

Qy	617	KMEHWSSFC	KRCRSQAVHLHLYGATYSAGEDRCACAGATHLLVQLRPETVLLDAYS	57
Db	551	KDSLH-----ELLHLH-----	-----DAMEKRLRFDDVTWNLSSYM	580
Qy	677	EHLAAALCTNP	LIETSLY-RNALGSRGVKLL--COGLRHFNPKCLQNLRLKRCRISSAC	733
Db	581	ESLSNALNQ	RSNKIQLDLWTHQQLNDDDVVYLAGCLG-----NISRLNMSHTYISSDQC	634
Qy	734	EDLSAALIANK	NLRMDLSGNGVGFPGMMLICEGLRHPCRL-----QMIIQR	781
Db	635	RVLKQALTE	QPSIQVHQH-----VPDILSTLVNVMRIIPDFNTSVYFVHDQOFESS	688
Qy	782	KC-----QLES	CACQEMASVLTGNPHLVELDTGNALEDLGLRLCOGLRHV--CRL	832
Db	689	KCWTSRGGK	LEVGGC-ELVVPVGALEKDVETIKLTAS-----LSLESEFLTPTLQCEL	741
Qy	833	RTLWLK-----	ICRLTAAC-----DELASTLS--V	856
Db	742	ASLTLYKKQV	TIKLTQHVLDKETIRRCVTLVYTRVTTVHWGKGLNHTDICSIIIDNII	801
Qy	857	NQSLRELDLS	NEIAGDGLV-----LLCGLRHPTCKQLTLRLGICRLGSA---ACEGLSV	909
Db	802	NKILDELKNA	FAKLMMKXVIITITILQDGVYEPSECVYINSLYQAGAVVWKVCWNLDL	861
Qy	910	V-----LQAN	NLRE--LDLSFNDL-----GDGCLWLLAEGLOHPAC--	944
Db	862	ATRENEQE	HEHFTQETWTPIUS-NDLIILCEKHETEAEVINIIPSGKIIPANOLNNSYCC	920
Qy	945	---RLQKLW	LDSCGLTAKA-C-----ENLYFTLIGINQTLTDLVLTNNALGD-TG	988
Db	921	NKKFKVVKQ	CTNEVHLIAKACGTFRWDDLWFPLTIITQKQOPVQVNIQIGNVTG	976
<p>RESULT 9</p> <p>T31668</p> <p>hypothetical protein COSI.5 - sea squirt (Ciona intestinalis)</p> <p>C:Species: Ciona intestinalis</p> <p>C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000</p> <p>C:Accession: T31668</p> <p>R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.</p> <p>submitted to the EMBL Data Library, October 1996</p> <p>A:Reference number: Z21050</p> <p>A:Accession: T31668</p> <p>A>Status: preliminary; translated from GB/EMBL/DBJ</p> <p>A:Molecule type: DNA</p> <p>A:Residues: 1-1075 <BTR></p> <p>A:Cross-references: EMBL:Z80904; PIDN:CAB02589.1</p> <p>C:Genetics:</p> <p>A;Introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2</p>				
Qy	102	YRDVYRRK	FLRMEDRNARLGEVCNLSHRYTRLLLVKEHNPMPVOOQLDTRGHARTVG	161
Db	2	HRD---	EKFVILQARAHATTGTVNEVFGSVQDVTHQHPEPAPVNVQL-----	47
Qy	162	HOASPIK	ITETPE-----PDERPPPR-----TVVMQGAAGICKSNLAHKVM	204
Db	48	-----VT	IDLQPLDKALENAQRYTEDESAFYKWIYERHANTVWVGPGPGVKTKTL--KMM	101
Qy	205	LDAWDGK	LFGQRFYDLFYINCREMOSATECSMODLI-----FSCWPEPSAPLQELIRV	258
Db	102	VQKILKH	ELPDTETIFFINVDIDFNKEMTLEFLTTNSRVKVNYTEESKALIFLHN	161
Qy	259	PERLLFI	DGDFELKPSFHDQPGMCLCWEKRPTELLNSLRKXLLPPELSLITRPT	318
Db	162	NENVAIF	FDGLDEASTNEFARIPHICKDGKSPVDIMKN-LFNLTLLPKAKIVVSTLH	220
Qy	319	ALEKHLRL	HEPRHVEILGFSAEKKEYFKYFHNAEQAGOVENVYR---DNEP-LFTMC	374

Db 221 QMYKLHPDYRPTSIPEVLGLLEAKNNL-----GTOLGCKYPAIKKILDDQPNLAHLC 274
 QY 375 FVPLVCWVCTCLOQLEGGLLRQTSRTTAYMYLILSLMQ-----PKGAPR 424
 Db 275 YLPINFILVFCILSN-EGSDI-----KTMTQVLFISMTFRFVLSHLKGEVPLDKVGAEM 328
 QY 425 LOPPPNQGLCSLAADGLWNOKILFEEQDLRKHGLDGE-----DVSFAFLNNIFQ 474
 Db 329 VK-----LARLAYKGLQORLKFETKTFDDVKLADWMTNFFHYTVDKSSGIRMKILE 381
 QY 475 KDINCERYSFTHLSQSFEEFAAMYYIILDEGEGAGDPQDVTRLLTYAFSERSFLAL--- 531
 Db 382 GN---KRSY-FTHLIWQEFVAAYVLM-----FVSYPEFEQLKPI 417
 QY 532 -----TSRFLGILLNEETRSLEKSLCWKVP-----HKMDLLOWIOSKASQDG 576
 Db 418 PKDAQWKRVRVGMFGICNPAYKQK-----LVFPATMIKDYEEKKELM--VPMMESLW 470
 QY 577 STLOQGSLEFFSCLYEIQBEETFOQALSHFQV---IWSNIAASKMEHMYSSFCLEKCRSA 633
 Db 471 SARGEDLIRFGWLHNEYNDDESKKPEDYLPVGLKMDAPKHLSEVKDLV--YALKSFTEP 528
 QY 634 QVHLH---YGATYSADGEDRARCAGAHLLV-----QLRP 666
 Db 529 HKRLDSYETTTTEVLETLRGVHTTTTITRFINNIEMKOSLMELLHLHLDAEELRF 588
 QY 667 ERTVLIDAYSEHLAALCTNPNIIEISLYRNAL-GSRGVKLL---COGLRHPNCKLQNLRL 723
 Db 589 IRVTNLSPMYERLSNAINQRSNKIQVLVKKHLHDDYDKYLAGCLG-----NISLLYM 642
 QY 724 KRCRISSACEDLSAAL 740
 Db 643 WGTDISSDQCSVLKQAI 659

RESULT 10
 A55478
 neuronal apoptosis inhibitory protein - human
 N:Alternate names: NAIP
 C:Species: Homo sapiens (man)
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
 C:Accession: A55478
 R:Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird, S.
 d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; Mackenzie, A.
 Cell 80, 167-178, 1995
 A:Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
 A:Reference number: A55478; MUID:95112344; PMID:7813013
 A:Accession: A55478
 A:Molecule type: mRNA
 A:Residues: 1-1232 <ROY>
 A:Cross-references: GB:U19251
 C:Genetics:
 A:Gene: GDB:SNAG; SNA
 A:Cross-references: GDB:120378; OMIM:600354; OMIM:253300
 A:Map position: 5q12.2-5q13
 C:Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prob
 F:94-110/Domain: transmembrane #status predicted <TM1>
 F:470-477/Region: nucleotide-binding motif A (P-loop)
 F:479-496/Domain: transmembrane #status predicted <TM2>
 F:476/Binding site: ATP (Lys) #status predicted
 F:618,923,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.4%; Score 183.5; DB 2; Length 1232;
 Best Local Similarity 21.3%; Pred. No. 1.3e-05;
 Matches 148; Conservative 94; Mismatches 248; Indels 205; Gaps 35;
 QY 188 MQGAGIGKSMIAHKVMDWADGKL-FQGRFYLYFYNCR-----EMN 229
 Db 468 VEGEAGSGKTVLLKKIAFLWASGCCPLNRFQVFLVLSLSTRPDEGLASIIQDLLEKE 527
 QY 230 QSATECSMODLFSWCWPEAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLWSE 289
 Db 528 GSVTECMENII-----QQL---KNQVFLDLDYKEI---CSIPO----- 561

QY 290 KRPTTELLNSLRKLLPELSLITRPTALEKHLLEHPRHVEILGFEARKEFYVK 349
 Db 562 -----VIQKLIQKNHLSRTCLLIARTNPARDIRYLE--TILEIQAP-----FYN 606
 QY 350 -----YFHNAEQAGQVFNVDNEPLFTMCFVPLVCWVCT-CLOQLEGGGLLQRT 400
 Db 607 TVCILRLKLSHNMTRLRKFMVYFGKNQSLQKIQKTFVAAICAHWFQPPD-----P 659
 QY 401 SRTTTAYMYLISL-MQPKGAPRLQPPNORGLCSLAADGLMNMOKILFEEODLRKHGL 459
 Db 660 SFDNVAVFKSYMERLSLRNKATAEILKATVSSCG--ELAKGFFSCCFEFNDLDAEAGV 717
 QY 460 DGEDVSAFLNMILFQKINCERYSFTHLSFOEFAAMYYI--LDRGEGAGDPQDV--- 514
 Db 718 DEDEDUTMCLMSKFTAQ-RLRPYRFLSFAFOEFLAGMLRIELDSDR---QSHQDLGLY 773
 QY 515 -----TRLITEYAF-----SERSFALTLSRFLGLNBEETRSHL----- 548
 Db 774 HLKQINSPMWTVSAYNNELNYVSSLPSTRAGPKIVSHLLHLVDNKESENISNDVYLKH 833
 QY 549 --EKSL-----CWKVSPIHKMDL-----LOWIOSKAQ 573
 Db 834 QPEISLQMLRLGLMQICQAYFSVMSEHLVLAKTAYQSNVTVAACSPFVLOFQ--- 889
 QY 574 SDGSTLOQGSLEFFSCLYEIQBEETFOQALSHFQV---IWSNIAASKMEHMYSSFCLEKCR 631
 Db 890 --GRITULGAL---NLQYFDPHESLULRSIHFS--IRGKTSPPRAHFSVLETCFDSQS 942
 QY 632 SAQVHLHYGATYS-----ADGED-----PARCSAGAHLLVLQRPRTVL--- 671
 Db 943 VPTIDQYASAFEPNMEWERNLAEKEDNVKSYMOMQRRASPDLSGTGKYLSPKQYKIPCL 1002
 QY 672 -----LDAYSEHLAALCT---NPNIIEISLYRNALGSRGVKLLCOGLRHPNCKLQNLRL 722
 Db 1003 EVDVNDIVVGQDMLMTVFSASQRIELHNLH---SRG---FIESIR-PALELSKAS 1054
 QY 723 LKRCRISSACEDLSAA---LIANKNLTRMDLSG 753
 Db 1055 VTKCSISKL---ELSAEQLLLTFLSLESLEVSIG 1086

RESULT 11
 B97746
 hypothetical protein RC0370 [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
 C:Accession: B97746
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
 Science 293, 2093-2098, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: B97746
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-312 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL02908.1; PID:gl5619435; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC0370

Query Match 3.3%; Score 180; DB 2; Length 312;
 Best Local Similarity 19.0%; Pred. No. 3.3e-06;
 Matches 64; Conservative 74; Mismatches 117; Indels 82; Gaps 12;
 QY 97 DQETRYDYVRKFRMLMEDNRNLGECVNLSHRY-----TRLLLVKEHSPNQ 144
 Db 34 DTEAKTREWFDKNYQVQEEKDNIAKSLNEENKFAILRRNLNTNKTSEIKNVEKTRNQ 93
 QY 145 VQQLLDTCGRGHARTVGHQASPIKTIETLPEPBERPEPRPVYVMOGAGIGKSMIAHKV 204
 Db 94 V-----DIK-----VLLGSAGIGKTLMLHYS 117
 QY 205 LQWADGKLFQGRFDYLYFYNCREMOSATE--CSMODLIFSCW-----PEPSAPLOEL--- 255

Db 118 YKGEKELNNKFYVPRIRLKKELLSWTVRGTTNIDDEILSCFVHYCLDSNDIKLEDIKS 177
 QY 256 IRVERLLFIIDGDELKPSFHDQGPWCLCWEKRPTELLNLSLRKLLPELSLIITT 315
 Db 178 IQDKRILLLLDGGDEV-----AFISQSNRRDRDIMDSVFOYK-----NIVWSS 221
 QY 316 RPTA-LEKHLRLHRLHPRHVEILGFSEARKEYFYKYFHNAOAGOVFNYYVDNE-PLFTM 373
 Db 222 RPNVIEEMSNRFB-QKVDNTGMDSEGIEQVHKNFYDKELGTPLKIFLDTHGQIKI 279
 QY 374 CFVPLVCWVCTCLOQLEGGLLRQTSRTTAYVML 410
 Db 280 CAVPINTALICLVWSQ-----AIRDFQKKTIVIKIL 311

RESULT 12

T52068
 RAN GTPase-activating protein 2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52068
 R:Pay, A.; Nick, P.; Nagy, F.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25929
 A:Accession: T52068
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-545 <PAY>
 A:Cross-references: EMBL:AF214560; PIDN:AAF25948.1

Query Match 3.3%; Score 179.5; DB 2; Length 545;
 Best Local Similarity 23.4%; Pred. No. 8.3e-06;
 Matches 71; Conservative 55; Mismatches 139; Indels 39; Gaps 6;
 QY 665 RPRTVL--LDAYSEHLAAALCTNPNLIELSYRNALGSRGVKLLCQGLRHPCNKLQNL 722
 Db 195 RPEALEVMNIFDALQGSLSLNSD-----NALGKGVRAFGALLKSLUS-SLEELY 248
 QY 723 LKRCRISSACEDLSAALIANKNLTRMDLSGN----- 754
 Db 249 LNDGIGSEKAAQAVSELIPSTENLRVLHPHNNMTGDEGALAAEVVKRSPLENFRCST 308
 QY 755 GVGFPQMLLCGLRHQP-CRLQMIQRLKQCSGACQEMASVLTGNPHLVLDLTGNAL 813
 Db 309 RVGSKGGIALSEALEH--CTHMEKLDLRNMFTEAGVSLSKTSSFKHMTLYLYLNL 366
 QY 814 EDGLRLCQGLRHPCRLRTLWLKICRLTAACDELASTLSVNQSLRELDLSLNLGDL 873
 Db 367 EDEGATAIVNALKESASPIEVLEMAGNDITVEASAIAACVAKQDLKLNLSNELKDE 426
 QY 874 GVLLLCGLRHPTCKLQTLRLGICRIGSAACGLSVYVQANHNRLRELDLSFNDLGDWGLW 933
 Db 427 GCQVIANCIIEVNSKLOQYIDMTNYYRAGARALAHVVVKKEAFKLNIDGNIISIEGIE 486
 QY 934 LLAE 937
 Db 487 ELKE 490

RESULT 13

T52063
 ran GTPase-activating protein [imported] - alfalfa
 C:Species: Medicago sativa (alfalfa)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52063
 R:Pay, A.; Nick, P.; Nagy, F.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25929
 A:Accession: T52063
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-533 <PAY>

A:Cross-references: EMBL:AF215731; PIDN:AAF19528.1
 A:Experimental source: subspecies x varia; cultivar A2

Query Match 3.3%; Score 178; DB 2; Length 533;
 Best Local Similarity 22.2%; Pred. No. 1e-05;
 Matches 112; Conservative 69; Mismatches 205; Indels 118; Gaps 16;

QY 510 PDQDVTLLIYAFSEERSFLATSLRFL-----GLNNE-----TRSHLE 549
 Db 19 PPSOSTRLM-----LVERWRNLTTPSIFSRKYGLNKDEAKEDAEAFVTSQHFE 74
 QY 550 K-----SLCWKVSPIHKMDLQWIOQSKAGSDSTLQOGLSFSCLYETQ---EEE 597
 Db 75 KEPDGDSSAVQIYAKESKMLLEVLRGPRGKEENGELISEKGDAAVETVEDISGGRRA 134
 QY 598 FI---QOALSHFQVIVVNSIAKMEHVMSSFCCLKRCRQAQ--VLHLYGATYSADGEDRARC 653
 Db 135 FIDQGEASELLKPLMGPNSTFKICFSNRSFGLDAAHVVEPMLISIKDQKEDVLSLSD----- 190
 QY 654 SAGATLLVQLRPRPTVL--LDAYSEHLAAALCTNPNLIELSYRNALGSRGVKLLCOGL 711
 Db 191 -----FIAGRPAEAIEVWNIIFSSALERAV-----LRYNLNSNANGKGVRAFRALL 238
 QY 712 RHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGVPQMMMLCGLRHP 771
 Db 239 KSNQ-LDEELYLMDNGISEEAAKAVAEILIPSTKLVLFHNNMTGDEGAFALAEVWKR 297
 QY 772 QCRLOMIQLRCKQLESAGACQEMASVLTGNPHLVLDLTGNA----- 812
 Db 298 PA-LEDFRCSSTRVSGEGVALAEALGACTHLKLDLRDNNMFGVAGVALSKVIPVFDL 356
 QY 813 -----LEDGLRLCQGLRHPCRLRTLWLKICRLTAACDELASTLSVNQSLREL 863
 Db 357 TEIYLSYNLEDDGAEALANALKEAPSELTLDMAGNITAKATYVSAECISSKQFLTKL 416
 QY 864 DLSLNLGDLGVLLLCGLRHPTCKLQTLRLGICRIGSAACGLSVYVQANHNRLRELDLS 923
 Db 417 NLSENELKDEGA-----GL-----ISKALEGRQGLSEVDLS 447
 QY 924 FNDLGDWGLWLLAEG-LQHPACRL 946
 Db 448 TNLITWSGAKLILAEAVVQKPGFKL 471

RESULT 14

A45841
 T-complex-associated-testes-expressed-1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
 C:Accession: A45841
 R:Garvetnick, N.; Tsai, J. Y.; Fox, H.; Pilder, S. H.; Silver, L.
 Immunogenetics 31, 283-284, 1990
 A:Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
 A:Reference number: A45841
 A:Accession: A45841
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-506 <SAR>
 A:Cross-references: GB:M28821; NID:G201909; PIDN:AAA40406.1; PID:G201910

Query Match 3.2%; Score 177.5; DB 2; Length 506;
 Best Local Similarity 21.6%; Pred. No. 1.1e-05;
 Matches 93; Conservative 55; Mismatches 155; Indels 127; Gaps 11;

QY 523 FERSFLATSLRPLFG-----LL-----NEETRSHLEKSLCWKVSPIHKMDLQWIO 569
 Db 152 FFERHLENLLKLFIPGTTDPNVLDLLPLCRNYYRIHVDDQL-----PVRMPTP--LQ 204
 QY 570 SKAQSDGSLTQOGLSFSCLYETQIEEIQOALSHFQVIVVNSIAKMEHVMSSFCCLKR 629
 Db 205 GEEQSDSGSEGEKS-----EPEKDIYQIQTVVGGKHLKHELDLVYGVKD 248
 QY 630 CRSAQVHLHYGATYSADGEDRARCAGATLLVQLRPERTVLLDAYSEHLAAALCTNPNL 699

Db 249 CGNFEWNLFLFY-RDCYSLAATIRACHTKI-FKLTRSKVDDKARILIRSLDHPAL 306
 QY 690 IELSLYRNALSGVKKLLCOGLRHPNCKLQNLRLKCRISSSACEDLSAALIANKNTRM 749
 Db 307 EELDLSHNLIGDRGARAANKLLSH----- 330
 QY 750 DLGNGVGFPWMMLLCEGLRHPQCRLOMIOLRKCQLESAGACQEMASVLTGTPHVELDLT 809
 Db 331 -----SRLRVNLANNQLOAPGAQSLAHALAHNTNVLNLR 367
 QY 810 GNALEDGLRLQLCOGLRHPVCRILTTLWTKICRLTAACDELASTLSVNSQSLRELDLSLNE 869
 Db 368 LNCIEDEGGQAIHALETNKC-LSVLHGGNKLSEPTATLLSQMLTWNTTIVSLNLSCHN 426
 QY 870 LGDLGVLLLCEGLRHPCTCKLTQRLGICRLGSAACEGLSVVLOAHNHLRELDLSFNDLGD 929
 Db 427 IGQDG-----GKOLLEGIS-----DNKTILEFDRLSDVSQ 457
 QY 930 WGLWLLAEGGL 939
 Db 458 ESEYLIGQVL 467

RESULT 15

T48193
 hypothetical protein F7A7.240 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48193
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: Z24487
 A:Accession: T48193
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-618 <BEV>
 A:Cross-references: EMBL:AL161946
 A:Experimental source: cultivar Columbia; BAC clone F7A7
 C:Genetics:
 A:Map position: 5
 A:Introns: 213/3; 255/3; 294/3; 350/1; 376/1; 419/2; 517/3; 601/3
 A:Note: F7A7.240

Query Match 3.2%; Score 172.5; DB 2; Length 618;
 Best Local Similarity 25.4%; Pred. No. 3.5e-05;
 Matches 122; Conservative 61; Mismatches 160; Indels 137; Gaps 28;
 QY 629 RCRAQVLHYGATYSADGEDRARCAGATHLLVQLRPBRTVLLDAYSEHLAAALCTNPN 688
 Db 123 KCVNLVEIDLSNATEMRDAD--AAVVAEARS-LEKLGRCKWLTDMGIGCIAVGCKLN 179
 QY 689 LIELSLYRNALSGVKKLL--COGLR-----HPNCKLQNLRL--LKRC-- 726
 Db 180 TVSLX-WCVGVGDGLGVLLAVKCKDIRTLDLSYLPITGKCLHDILKQHLLEILLGECFG 238
 QY 727 -----RISSACEDLSAALIANKNLTMDLSGNGVGFPGPMWLLC 765
 Db 239 VDDDSLKSLRHDCKSLKMYKQKLDASSCQNLIT-----HRLTSL-LSGAGY----- 283
 QY 766 EGLRHPQCKLQMIOLRKCQLESAGACQEMASVLTGTPHVELDLTGNALDELGLR---LLC 822
 Db 284 -----LQRLDLSHC--SSVISLDFASSLKKVKSALQSIKRLDGCSTPDLKAIGTLC 332
 QY 823 QGLRH-----PVCRRLTLM--KICRLTAACDELASTLSVNSQSLRELDLSLNEGLDL 873
 Db 333 NSLKEVSLKCVSVTDEEAFWLGQRCRL-----LEELDTDNEIDD- 374
 QY 874 GVLLECEGLRH-PTC-KLQTLRLGICRLGSAACEGLSVVLOAHNHLRELDLSFNDIG--D 929
 Db 375 -----EGLKSISSCULSSLSKLGICL--NITDKGLSYIGMGCNSLRELDL-YRVSIGITD 425

QY 930 WGLWLLAEGLOHPACRLQKWLWDSC-GLTAKACENLYFTLGINQTLTDL---YLTNNALG 985
 Db 426 VGISTIAQGCIIH-----LETINISYCQDITDKSLVSL-SKCSLLQTFESRCPNITSGGLA 480
 QY 986 DTGVRLLCKRLSHPGCK-----LRVLWLFQMDLNKMTSHSLAALRVTKPYL-DIGC 1035
 Db 481 AIAVR--CKRLAKVDLKKCPISINDAGLLALAHFSONLKQINVSDTAVTEVGLLSLANIGC 538

Search completed: July 30, 2004, 13:49:54
 Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:43:01 ; Search time 18 Seconds
(without alignments)
2994.034 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472

Sequence: 1 MLRTAGDGLCLSTYLSL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5433	99.3	1061	1 NAL2 HUMAN	P59046 homo sapien
2	2396	43.8	1034	1 C1S1 HUMAN	Q96P20 homo sapien
3	2359.5	43.1	1033	1 C1S1 MOUSE	Q84b8 mus musculus
4	1799	32.9	1093	1 NAL4 HUMAN	Q86w24 homo sapien
5	1457.5	26.6	994	1 NAL4 HUMAN	Q96m22 homo sapien
6	1375	25.1	1473	1 NAL5 HUMAN	Q9C000 homo sapien
7	1320	24.1	1200	1 NAL5 HUMAN	P59047 homo sapien
8	1309	23.9	1062	1 NAL2 HUMAN	Q9nx02 homo sapien
9	1163	21.3	980	1 NAL7 HUMAN	Q8wx94 homo sapien
10	1160.5	21.2	1111	1 NAL5 MOUSE	Q9rlm5 mus musculus
11	1159	21.2	892	1 NAL6 HUMAN	P59044 homo sapien
12	1115	20.4	843	1 NAL6 MOUSE	Q91ws2 mus musculus
13	1078.5	19.7	1033	1 NAL1 HUMAN	P59045 homo sapien
14	1067	19.5	854	1 NAL6 RAT	Q63035 rattus norv
15	724	13.2	460	1 RINI HUMAN	P13489 homo sapien
16	720	13.2	460	1 RINI PANTR	Q8h2p9 pan troglod
17	717	13.1	456	1 RINI RAT	P29315 rattus norv
18	709	13.0	456	1 RINI PIG	P10775 sus scrofa
19	529	9.7	1040	1 CARP HUMAN	Q9hc29 homo sapien
20	522	9.5	953	1 CAR4 HUMAN	Q9y239 homo sapien
21	519	9.5	1020	1 CARP MOUSE	Q8k3z0 mus musculus
22	469.5	8.6	953	1 CAR4 MOUSE	Q8bbh0 mus musculus
23	377.5	6.9	1130	1 C2TA HUMAN	P39076 homo sapien
24	317.5	5.8	1155	1 C2TA MOUSE	P79621 mus musculus
25	260	4.8	1024	1 BIRC HUMAN	Q9np44 homo sapien
26	200.5	3.7	1403	1 BIRC MOUSE	Q9r016 mus musculus
27	197	3.6	1403	1 BIR4 MOUSE	Q9gwk5 mus musculus
28	189	3.5	1403	1 BIR1 HUMAN	Q13075 homo sapien
29	182.5	3.3	1403	1 BIRF MOUSE	Q9jib6 mus musculus
30	179.5	3.3	1402	1 BIRG MOUSE	Q9jib3 mus musculus
31	174.5	3.2	596	1 RGPI DROME	Q9v1w3 drosophila
32	172.5	3.2	580	1 RGPI XENLA	Q13066 xenopus lae
33	171	3.1	1447	1 BIRB MOUSE	Q9qk4 mus musculus

RESULT 1					
NAL2 HUMAN					
ID	NAL2_HUMAN	STANDARD;	PRT;	1061	AA.
AC	P59046;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	10-OCT-2003 (Rel. 42, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	NACHT-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like protein 7) (Monarch-1).				
DE	like protein 7) (Monarch-1).				
GN	NALP12 OR PYPAF7.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=22162427; PubMed=12019269;				
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,				
RA	Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;				
RT	"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates				
RT	activation of NF-kappa B and caspase-1-dependent cytokine				
RT	processing.";				
RL	J. Biol. Chem. 277:29874-29880(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22451042; PubMed=12563287;				
RA	Tschopp J., Martinon F., Burns K.;				
RT	"NALPs: a novel protein family involved in inflammation.";				
RL	Nat. Rev. Mol. Cell Biol. 4:95-104(2003).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).				
RC	TISSUE=Lymphoma;				
RA	Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;				
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Leukocyte;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,				
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length				

ALIGNMENTS

34	161.5	3.0	589	1	RGPI_MOUSE	P46061 mus musculus
35	159	2.9	1378	1	WR52_ARATH	Q9f883 arabidopsis
36	155	2.8	587	1	RGPI_HUMAN	P46060 homo sapien
37	152.5	2.8	2493	1	CYAA_USTMA	P46060 ustilago ma
38	139	2.5	578	1	LR15_RAT	Q8r5m3 rattus norv
39	136.5	2.5	767	1	MEPV_MOUSE	Q9j126 mus musculus
40	134.5	2.5	906	1	RD14_ARATH	Q9x1f0 arabidopsis
41	132.5	2.4	907	1	LGR5_HUMAN	Q75473 homo sapien
42	132	2.4	907	1	LGR5_MOUSE	Q9z1p4 mus musculus
43	131.5	2.4	1256	1	FLI1_DROME	Q24020 drosophila
44	130	2.4	1138	1	RD13_ARATH	Q8w3k0 arabidopsis
45	129	2.4	463	1	POF2_SCHPO	O74783 schizosacch

RT human and mouse cDNA sequences.":

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: May mediate activation of CASPI via ASC and promote

CC activation of NF-kappa-B via IKK.

CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Name=1; Synonyms=1;

CC IsoId=P59046-1; Sequence=Displayed;

CC Name=2; Synonyms=11;

CC IsoId=P59046-2; Sequence=VSP_005524;

CC Name=3; Synonyms=111;

CC IsoId=P59046-3; Sequence=VSP_005523;

CC -!- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,

CC predominantly in eosinophils and granulocytes, and at lower levels

CC in monocytes.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AY095146; AAM18227.1; -

CC EMBL; AY154467; AAO18163.1; -

CC EMBL; AY116204; AAM75142.1; -

CC EMBL; AY116205; AAM75143.1; -

CC EMBL; AY116206; AAM75144.1; -

CC EMBL; BC028069; AAM75145.1; -

CC Genew; HGNC:22938; NALP12.

CC InterPro; IPR001611; LRR.

CC InterPro; IPR007091; LRR_Rninh.

CC InterPro; IPR003590; LRR_Rninh_sub.

CC InterPro; IPR007111; NACHT_NTPase.

CC InterPro; IPR004020; PAAD_DAPIN_dom.

CC Pfam; PF00560; LRR; 2.

CC Pfam; PF02758; PAAD_DAPIN; 1.

CC SMART; SM00368; LRR_RT; 11.

CC PROSITE; PS00824; DAPIN; 1.

CC PROSITE; PS00837; NACHT; 1.

CC ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.

KW

FT DOMAIN 1 95 DAPIN.

FT REPEAT 211 528 NACHT.

FT REPEAT 712 735 LRR 1.

FT REPEAT 741 764 LRR 2.

FT REPEAT 769 791 LRR 3.

FT REPEAT 798 821 LRR 4.

FT REPEAT 826 849 LRR 5.

FT REPEAT 883 906 LRR 6.

FT REPEAT 940 968 LRR 7.

FT REPEAT 997 1020 LRR 8.

FT NP_BIND 217 224 ATP (POTENTIAL).

FT VARSPLIC 862 973 Missing (in isoform 3).

FT VARSPLIC 976 1031 /FTId=VSP_005523.

FT CONFLICT 691 691 L->LR (IN REF. 3).

FT SEQUENCE 1061 AA; 120172 MW; 8C10AFE4907C131B CRC64;

Query Match 99.3%; Score 5433; DB 1; Length 1061;

Best Local Similarity 97.4%; Pred. No. 0;

Matches 1034; Conservative 0; Mismatches 0; Indels 28; Gaps 2;

QY 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLGTATLGRGKIPWGSMEKAGPLEMAQLL 60

DB 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLGTATLGRGKIPWGSMEKAGPLEMAQLL 60

QY 61 ITHFGPEAWRLALSTFFRINRKLWERGQREDLVYR----- 96

DB 61 ITHFGPEAWRLALSTFFRINRKLWERGQREDLVYR----- 120

QY 97 ---DPQETVYRVRRKFRIMEDRNARLGEVNLSHRYTRLLLVKHSNPMQVQOQLDYG 153

DB 121 PRKDPQETVYRVRRKFRIMEDRNARLGEVNLSHRYTRLLLVKHSNPMQVQOQLDYG 180

QY 154 RGHARTVGHQASPIKETLFEDEPERPEPRVVMQGAAGIGKSLAHKVLMDWDGKLF 213

DB 181 RGHARTVGHQASPIKETLFEDEPERPEPRVVMQGAAGIGKSLAHKVLMDWDGKLF 240

QY 214 QGRFDYLFYINCREMNQSMQDLIFSCWPEPSAPLQELIRVPERLLFIIDGDELK 273

DB 241 QGRFDYLFYINCREMNQSMQDLIFSCWPEPSAPLQELIRVPERLLFIIDGDELK 300

QY 274 PSFHPDQGPWCWCWEKRPTELLNSLIRKLLPELSLITTRPTALEKHLLEHPRV 333

DB 301 PSFHPDQGPWCWCWEKRPTELLNSLIRKLLPELSLITTRPTALEKHLLEHPRV 360

QY 334 EILGFSEAEKRYFYKYFHNAEQGVFNYVRDNEPLFTMCVFPLVVCVVCTCLOQLLEG 393

DB 361 EILGFSEAEKRYFYKYFHNAEQGVFNYVRDNEPLFTMCVFPLVVCVVCTCLOQLLEG 420

QY 394 GGLLRQTSRTTAVVYMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEED 453

DB 421 GGLLRQTSRTTAVVYMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEED 480

QY 454 LRKHGLDGDVSAFLNMNIFQKINCERYYSPTILSPQEFFFAAMYVILDEGEGAGDQD 513

DB 481 LRKHGLDGDVSAFLNMNIFQKINCERYYSPTILSPQEFFFAAMYVILDEGEGAGDQD 540

QY 514 VTRLLTEYAFSERSPLATSRFLGLNEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQ 573

DB 541 VTRLLTEYAFSERSPLATSRFLGLNEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQ 600

QY 574 SDGSTLQGSLEFFSCLYEIQEETIQOALSHFQVIVNSIASKMEHVVSSFCCLKRCSA 633

DB 601 SDGSTLQGSLEFFSCLYEIQEETIQOALSHFQVIVNSIASKMEHVVSSFCCLKRCSA 660

QY 634 QVHLHYGATYSADGEDRARCAGAHLLIQLRPERTVLLDAYSEHLAAALCTPNLIELS 693

DB 661 QVHLHYGATYSADGEDRARCAGAHLLIQLRPERTVLLDAYSEHLAAALCTPNLIELS 719

QY 694 LYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 753

DB 720 LYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 779

QY 754 NGVGFPMMLLCEGLRHPQCRLOMIQLRKQLESGACQEMASVLTGNPHLVLDLTGNAL 813

DB 780 NGVGFPMMLLCEGLRHPQCRLOMIQLRKQLESGACQEMASVLTGNPHLVLDLTGNAL 839

QY 814 EDLGLRLCOGLRHPVCRILTMLKICELTAACDELASTLSVNQSLRELDLSNELGDL 873

DB 840 EDLGLRLCOGLRHPVCRILTMLKICELTAACDELASTLSVNQSLRELDLSNELGDL 899

QY 874 GVLLICEGLRHPCKLQTLRLGI CRLSAAACEGLSVVQLAHNHLRELDLSFNDLGDWGLW 933

DB 900 GVLLICEGLRHPCKLQTLRLGI CRLSAAACEGLSVVQLAHNHLRELDLSFNDLGDWGLW 959

QY 934 LLAEGLQHPACRLQKWLDCGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVLLC 993

DB 960 LLAEGLQHPACRLQKWLDCGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVLLC 1019

QY 994 KRLSHPCCKLRVLWFGMDLNKMTSHSLAALRVTKPYLDIGC 1035

DB 1020 KRLSHPCCKLRVLWFGMDLNKMTSHSLAALRVTKPYLDIGC 1061

RESULT 2

CIS1 HUMAN

ID -CIS1 HUMAN STANDARD; Q8TEU9; Q8WXH9; PRT: 1034 AA.

AC Q96P20; Q75434; Q8TCW0; Q8TEU9; Q8WXH9;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cold autoimmune syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
 DE PYD-containing protein 3) (PYRIN-containing APAF1-like protein 1)
 DE (Angiotensin/vasopressin receptor AII/AVP-like).
 GN CIAS1 OR NALP3 OR PYPAF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS FCAS MET-198; VAL-439
 RP AND GLT-627, AND VARIANT MWS VAL-352.
 RX MEDLINE=21547523; PubMed=11687797;
 RA Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A.,
 RA Kolodner R.D.;
 RT Mutation of a new gene encoding a putative pyrin-like protein causes
 RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome.";
 RN Nat. Genet. 29:301-305(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RP MEDLINE=21909508; PubMed=11786556;
 RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
 RA Mak S., Lora J.M., Briskin M., Jurman M., Cao J., DiStefano P.S.,
 RA Bertin J.;
 RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
 RT and activates NF-kB";
 RN J. Biol. Chem. 277:11570-11575(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), VARIANT MWS MET-198, AND
 RP VARIANTS FCAS/MWS TRP-260 AND PRO-305.
 RX MEDLINE=22241334; PubMed=12355493;
 RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
 RA Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
 RA Cotter F.E., Thome M., Hitman G.A., Tschopp J., McDermott M.F.;
 RT "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad
 RT phenotype including recurrent fever, cold sensitivity, sensorineural
 RT deafness, and AA amyloidosis.";
 RN Arthritis Rheum. 46:2445-2452(2002).
 RN [4]
 RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
 RC TISSUE=Blood;
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.;
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 RT 300 previously undefined genes expressed in CD34+ hematopoietic
 RT stem/progenitor cells.";
 RN Genome Res. 10:1546-1560(2000).
 RN [5]
 RP VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-349; THR-439 AND
 RP ARG-569, AND VARIANT FCAS/MWS TRP-260.
 RX MEDLINE=21987640; PubMed=11992256;
 RA Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
 RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,
 RA Granel B., Frances C., Garcier F., Ederly P., Boulinguez S.,
 RA Domergues J.-P., Delpech M., Grateau G.;
 RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
 RT and familial cold urticaria: a novel mutation underlies both
 RT syndromes.";
 RN Am. J. Hum. Genet. 70:1498-1506(2002).
 RN [6]
 RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
 RP THR-662, AND TISSUE SPECIFICITY.
 RX MEDLINE=22062556; PubMed=12032915;
 RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
 RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
 RT "Chronic infantile neurological cutaneous and articular syndrome is
 RT caused by mutations in CIAS1, a gene highly expressed in
 RT polymorphonuclear cells and chondrocytes.";
 RN Am. J. Hum. Genet. 71:198-203(2002).
 RL

CC -!- FUNCTION: May function as a potential inducer of apoptosis.
 CC Interacts selectively with apoptosis-associated specklike protein
 CC containing a CARD domain (ASC). This complex may function as an
 CC upstream activator of NF-kappaB signaling.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=2;
 CC IsoId=Q96P20-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;
 CC Name=3;
 CC IsoId=Q96P20-3; Sequence=VSP_005519;
 CC -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly
 CC expressed in polymorphonuclear cells, undetectable or expressed
 CC at a lower magnitude in B and T lymphoblasts, respectively. High
 CC level of expression detected in chondrocytes. Low or no expression
 CC in the other tissues tested.
 CC -!- DISEASE: Defects in CIAS1 are a cause of familial cold
 CC autoinflammatory syndrome (FCAS) [MIM:120100]; commonly known as
 CC familial cold urticaria. FCAS is rare autosomal dominant systemic
 CC inflammatory disease characterized by episodes of rash,
 CC arthralgia, fever and conjunctivitis after generalized exposure to
 CC cold.
 CC -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome
 CC (MWS) [MIM:191900]; a rare autosomal dominant fever syndrome with
 CC episodic urticaria, arthralgia, amyloidosis and progressive
 CC sensorineural deafness.
 CC -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile
 CC neurologic cutaneous and articular syndrome (CINCA) [MIM:607115];
 CC also known as 'neonatal onset multisystem inflammatory disease,'
 CC or NOMID, a rare congenital inflammatory disorder characterized by
 CC a triad of neonatal onset of cutaneous symptoms, chronic
 CC meningitis, and joint manifestations with recurrent fever and
 CC inflammation.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts
 CC in positions 893, 918 and 926.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AF410477; AAL33908.1; -;
 CC EMBL; AF427617; AAL33911.1; -;
 CC EMBL; AY051117; AAL12497.1; -;
 CC EMBL; AY051112; AAL12497.1; JOINED.
 CC EMBL; AY051113; AAL12497.1; JOINED.
 CC EMBL; AY051114; AAL12497.1; JOINED.
 CC EMBL; AY051115; AAL12497.1; JOINED.
 CC EMBL; AY051116; AAL12497.1; JOINED.
 CC EMBL; AY056059; AAL12497.1; JOINED.
 CC EMBL; AY056060; AAL12497.1; JOINED.
 CC EMBL; AY051117; AAL12498.1; -;
 CC EMBL; AY051112; AAL12498.1; JOINED.
 CC EMBL; AY051113; AAL12498.1; JOINED.
 CC EMBL; AY051114; AAL12498.1; JOINED.
 CC EMBL; AY051115; AAL12498.1; JOINED.
 CC EMBL; AY051116; AAL12498.1; JOINED.
 CC EMBL; AF420469; AAL65136.1; -;
 CC EMBL; AF468522; AAL78632.1; ALT_INIT.
 CC EMBL; AF092033; AAM14669.1; ALT_INIT.
 CC EMBL; AF418985; AAL14640.2; ALT_INIT.
 CC EMBL; AF054176; AAC39910.1; ALT_FRAME.
 CC EMBL; HGNC:16400; CIAS1.
 CC MIM; 606416; -;
 CC MIM; 120100; -;
 CC MIM; 191900; -;
 CC

[illegible]

CC CC -1- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
 CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 DR EMBL; AV154469; AAO18165.1; --
 DR Genew; HGNC:22939; NALP14.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR_2; DAPIN.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00368; LRR_R1; 12.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.
 KW DOMAIN
 FT 1 97
 FT DOMAIN
 FT 177 499
 FT REPEAT
 FT 525 550
 FT REPEAT
 FT 672 695
 FT REPEAT
 FT 733 756
 FT REPEAT
 FT 757 780
 FT REPEAT
 FT 785 807
 FT REPEAT
 FT 842 869
 FT REPEAT
 FT 871 894
 FT REPEAT
 FT 899 926
 FT REPEAT
 FT 956 979
 FT REPEAT
 FT 985 1008
 FT REPEAT
 FT 1013 1037
 FT REPEAT
 SQ SEQUENCE 1093 AA; 124732 MW; 124EEACE22A11D6F CRC64;
 Query Match 32.9%; Score 1799; DB 1; Length 1093;
 Best Local Similarity 37.4%; Pred. No. 2.8e-122;
 Matches 413; Conservative 177; Mismatches 388; Indels 126; Gaps 15;
 QY 13 LSTYLEELAEVLEKFKLYLGATGTELGKIPGMSMEKAGPLEMAQLLIHFPGEAWRL 72
 DB 15 LLLYLEELNKEELNFTFLKKTETMEPEHGLTPWNEVKKARREDLANLMKKYFGEKAVS 74
 QY 73 ALSTPERINRKLWERGORE-----DLVRDPQET-----YRDYVRRKF 110
 DB 75 SLKIFGMLKDLCEAKEINWSAQTIGPDDAKAGETQEDQEAVLGDGTENRIKEKF 134
 QY 111 LKMDERNARLGECVNI-SHRYTRLLLVKHSNPNVQVQQLDTRGRGHARTVGHQASPIKE 170
 DB 135 CITWDRKSLAGKPEDFHG-----IAEKDRKL-----LE 163
 QY 171 TLEPDEPERPEPTVMQGAAGIGKMLAHKVLMDWADGKLFQGRDYLYFVINCENMQ 230
 DB 164 HLFEDVDVKTGAQPIVQLQGAAGVKTTLVRKAMLDWAGSLYQQRKYFYFLNGREINQ 223
 QY 231 SATCESMQDLIFSCWPEPSAPLOELIRVPERLLFIIDGFDLPKSPHDPQPGWCLQWEEK 290
 DB 224 -LKERSFAQLISKDWSSTEGPIEEIMVQPSLLFIIDSPDELNFAPEEPFALCEDWTQE 282
 QY 291 RPTELLNLSIRKKLPELPELSLLITTPRTALEKHLRLLHPRHVRVILGFSAEKREYFYKY 350
 DB 283 HPVSFLMSSLRKMVLPASLLVTRITTSKRLKQLLNHHYVELLGMSEDAEEYIYQF 342
 QY 351 FHNAEQAGQVFNVRDNEPLFTWCFFPLVCMVCTCLQQLEGGLLRQTSRTTAYVYL 410
 DB 343 FEDRWAMKVFSSIKSNEMLFSMCQVPLVCWAACCTCLKQMEKGGDVTLCQYTTALFTC 402
 QY 411 YLLSLMQP-KPGAPRLQPPNPQ---RGLCSLAADGLWNOKILFEEQDLRKHGLDGEDVSA 466

DB 403 YISLFTPDGGSPL---PNOAQLRRLCOVAAGKWTMTYVYFRENLRRLGLTQSDVSS 459
 QY 467 FLNNMIFOKDINCERYYSFIHLSFQFFAAMYIL-----DEGEGAGPDQDVTRLLTEVA 522
 DB 460 FMDSNIIQKDAEYENCYVFTHLVQVEFFAAMFYMLKGSWEAGNPSQPPEDLSLLQSTS 519
 QY 523 FSRFSFLATLSFLGLLNEETRSHLEKSLCKWVSPHMKMDLLOMTQSKAQSDGSLQOQ 582
 DB 520 YKD-PHLTQMKCFLEGLLNEDRVKQLERTFNCKMSLKIKSKLLQCMVEVLGNDYSPSOLG 578
 QY 583 SLEFFSCLYIEQEBEFOQALSHFQVIVVSNTASKMEHVMVSFCLKRCRQAQVHLHYGA- 641
 DB 579 FLELFHCLYETQDKAFISQAMCFPKVAI-NICEKHLLVSSFCCLKRCCLRTIRLSVTV 637
 QY 642 -----TYSADGEDRARC-----SAGHTLLVQLRPERTVLLDAYSEHL----- 679
 DB 638 VFEKKILKTSLPTNTWDGRIITHCWDLCSLVLT-----NEHLRELDYHNSLDKSAMN 691
 QY 680 -----AAALCTNPNLIELSLYRNALGSRGVKLLCQ 709
 DB 692 ILHHLRHPNCKLQKLLKFIPTFDGCGQDISTSILHNKLMHLLDKSGDIGNGVKSLCE 751
 QY 710 GLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTMOLDSGNGVGFPGMMLCEGLR 769
 DB 752 ALKHPECKLQTLRLSCNLTVPCCNLISNALIRSQSLIFNLSTNNLLDDGVQLLCEALR 811
 QY 770 HPQCRLOMIQLRCKQLESCAGCEMASVLTGNPHLVELDITGNALEDLGLRLLCQGRHPV 829
 DB 812 HPKCYLERLSLESCGLEAGCEYLSIALISNKRHLTHCLADNVLDGGGVKJMSDALQHAQ 871
 QY 830 CLRRLTWLKI CLRATAACDELASTLSVNOSLRELDLSLNLGDLGVLLLCCEGLRHTCKL 889
 DB 872 CTLKSLVLRCHFTSLSSSEYLSLHNSLTHLDLSGNLQDNGVKLLCDVFRHSCNL 931
 QY 890 QTLRLGICRLGSAACBGLSVLQANHLRELDLSPNDLGDWGLWLLAEGLQHPACLQKL 949
 DB 932 QDLMLGCVLTWACCLDLASVILNPNRLSLDGNNDLQDDGVKILCDALRYPCNCIQL 991
 QY 950 WLDSCLGTAKACENLYFTLGINQTLTDLVLTNNALGDTGVRLLCRKLSPHGGCKRLVWLWF 1009
 DB 992 GLEYCGLTSLCCQDSSALICNKRLLKMLNTQNTLGYEGIVKLYKVKSPKCKLQVLGLC 1051
 QY 1010 GMDLNKMTSRLAALRVTKPYLDI 1033
 DB 1052 KEAFDEAEAKLLEAVGVSNPHLII 1075
 RESULT 5
 ID NAL4 HUMAN STANDARD; PRT; 994 AA.
 AC Q96MK2; Q96AY6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 4 (PYRIN-containing APAF1-like
 DE protein 4) (PAAD and NACHT-containing protein 2) (Ribonuclease
 DE inhibitor 2).
 GN NALP4 OR PYPAF4 OR PAN2 OR RNH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RK MEDLINE=22451042; PubMed=12563287;
 RA Tschopp J., Martinon F., Burns K.;
 RT "NALP4: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP Fiorentino L., Reed J.C.;
 RT "Pan2, a novel PAAD-containing protein.";

DR EMBL; AF229062; AAK00751.1; --
 DR EMBL; AB023143; BAA76770.1; --
 DR EMBL; AL117470; CAB55945.1; --
 DR PIR; T17255; T17255.
 DR HSSP; P13489; 144Y.
 DR MIM; 606636; --
 DR GO; GO:0005622; C:intracellular; IC.
 DR GO; GO:0016506; F:apoptosis activator activity; NAS.
 DR GO; GO:0008656; F:caspace activator activity; NAS.
 DR GO; GO:0019899; F:enzyme binding; IPI.
 DR GO; GO:0006919; P:caspace activation; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; NAS.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02759; PAAD_DAPIN; 1.
 DR PRINTS; PR00364; DISEASERISIT.
 DR PROSITE; PS02029; CARD; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 92 DAPIN.
 FT DOMAIN 328 637 NACHT.
 FT REPEAT 704 725 LRR 1.
 FT REPEAT 807 830 LRR 2.
 FT REPEAT 864 887 LRR 3.
 FT REPEAT 921 944 LRR 4.
 FT REPEAT 950 973 LRR 5.
 FT REPEAT 1199 1215 LRR 6.
 FT REPEAT 1216 1236 LRR 7.
 FT DOMAIN 1374 1463 CARD.
 FT NP_BIND 334 341 ATP (POTENTIAL).
 FT VARSPLIC 958 987 Missing (in isoform 3 and isoform 4).
 FT VARSPLIC 1262 1305 /FTid=VSP_004326.
 FT VARSPLIC 340 340 Missing (in isoform 2 and isoform 3).
 FT MUTAGEN 340 340 /FTid=VSP_004327.
 FT MUTAGEN 340 340 K->L: NO EFFECT.
 FT CONFLICT 155 155 L -> H (IN REF. 1).
 FT CONFLICT 246 246 T -> S (IN REF. 1).
 FT CONFLICT 782 782 T -> S (IN REF. 1).
 FT CONFLICT 878 878 T -> M (IN REF. 1).
 FT CONFLICT 995 995 T -> I (IN REF. 1).
 FT CONFLICT 1119 1119 M -> V (IN REF. 1).
 FT CONFLICT 1184 1184 M -> V (IN REF. 1 AND 6).
 FT CONFLICT 1241 1241 V -> L (IN REF. 1).
 FT CONFLICT 1366 1366 R -> C (IN REF. 1).
 SQ SEQUENCE 1473 AA; 438F0DCE45C2562D CRC64;
 Query Match 25.1%; Score 1375; DB 1; Length 1473;
 Best Local Similarity 33.9%; Pred. No. 2.2e-91;
 Matches 366; Conservative 138; Mismatches 357; Indels 218; Gaps 23;
 QY 12 RLSTVLELEAVELEKFKKFLYLTA--TELGGKIPWGSMEKAGPLEMAQLLITHFGPEEA 69
 DB 8 RLACYLEFLKKEELKEFOLLANKAHSRSSGETP-AQPEKTSMEVASLYVAQYGEORA 66
 QY 70 WRALSTFERINRKDLWERGQR-----EDLVRDPQETRYDYV----- 106
 DB 67 WDLALHTWEQMLRSLCAQAQAGAGCHSPSPFPYSPPEHLGSPSQPTSTAVLMPWIHELPA 126
 QY 107 -----RRKFLMEDRNALGECVNLSHRYTRLLLVKEHNPQVQ----- 146
 DB 127 GCTQGSERVLQLPDTSGRWRREISASLLYQALFSPDHPSPSPESNAPTSTAVLGSW 186
 QY 147 ----- 146

Db 187 GSPQPSLAPREQEAPGTQWPLDSTSGIYYTEIREREREKSEKGRPPWAAVVGTPPQAHT 246
 QY 147 -----QQLDTGRGHARTV----- 160
 Db 247 SLQPHHPWPEPSVRESLCSTWPKNEDFNQKFTQLLLQRPSPRSQDPLVKRSWPDYVEE 306
 QY 161 --GQASPIKIEITLFEDEERPEPRVVMGAGICKSLAHKVMLDWADGKLPQGRD 218
 Db 307 NRHGL---IEIRDLFGGLDTQE--PRVILQGAAGIKSTLARQVKEAWGRGQYIGRQFQ 362
 QY 219 YLFYINCKRMNQSAATECSMODLIFSCWPSPAPIQELIRVPERLFIIDGFDLKPFSHD 278
 Db 363 HVFYFSCRELAQSKV-VSLAELIGKDTATPAPIRQILSRPERLLFILDGVDFGWVLOE 421
 QY 279 PQGPMCLCWEEKRPTTELLNSLIRKLLPELSLITRPTALEKHLRLLLEHPRHVEILGF 338
 Db 422 PSELCLHWSQPQPADALLGSLGKTLIPASFLITARTALQNLIPSLEQARVVEVLGF 481
 QY 339 SEAEKKEVYKYFHNAEQAGOVNVRDNEPLFTMCVPLVVCVWCTCLOQLEGGLLR 398
 Db 482 SSSRKEYFYRYFTDERQAIRAFRLVSKNELWALCLVPPVWSLACTCLMQMKRKEKLT 541
 QY 399 QTSRTTAVVMLYLLSLMQPKGAPRLQPPNQREGCLSLAADGLWNQKILFEEODLRKHG 458
 Db 542 LTSKTTTTLCHYLAQALQAQPLGPQL-----RDLCSLAAEGIQWKTLPSPDLRKHG 595
 QY 459 LDGEDVSALNMNIPQKINCERYYSFIHLSFQFFFAAMYIILDEGEGAGPDQ----DV 514
 Db 596 LDGAIISTFLKMGILQEH-PIPLSYSTHLCFQFFFAAMSYVL-EDEKGRKHSNCIIDL 653
 QY 515 TRLLTEYAFSERSFALTSRELFGLNNEETSHLEKSLCMKVSPIHKMDLLQWTSQAQS 574
 Db 654 EKTLEAYGI-HGLFGASTTRFLLGLLSDGEREMENIFHCRLSQ--GRNLMQWVPSLO-- 708
 QY 575 DGSTLQOQSLEFFSCLYEIOEEETIQALSHFQIVIVSNIASKMEHVMSSFLKRCRQAQ 634
 Db 709 --LLLQPHSLSLCHLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK 765
 QY 635 VLHLYGATYSADGEDRARCASGAHTLLVQLRPETVLLDAYSEHLAAALCTNPNLIELSL 694
 Db 766 KLQLI-----EGRQHRSTWSPTMVFLR--VWPVTDAYWQILFSLVKVTRNLKELDL 815
 QY 695 YRNALSGRVKLLCQGLRHNPCKLQNLRLKRCRISSSACEDLSAALIANKNLTMDLSGN 754
 Db 816 SGNLSHSAVSKLCKTLRRPCLLETURLACGTAEDCKDLAFGLRANQITELDLDFN 875
 QY 755 GVGFPGMMLLCEGLRHPQCRLOMIQLRKKCQLESACQEMASVGLTNPHLVELDTGNALE 814
 Db 876 VLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNLD 935
 QY 815 DLGLRLCQGLRHPVCRILRTLWLKICRLTAACDELASTLSVNQSLRELDLSINEL---- 870
 Db 936 DVGVRLLCEGLRHPACKLIRLGL-----DQTLSDDEMQLRALQEQKPOLLIIFS 985
 QY 871 -GDGLVILLCEGL-----RHPTCKIQLTLRLGICRLGSAACEGLSVLQANHLRELDLS 923
 Db 986 RRFQSVMTPEGLDTGEMSNSTSLKRLQRLGSEAAHSV-----AQANLKLLDVS 1035
 RESULT 7
 ID NAL5 HUMAN STANDARD; PRT; 1200 AA.
 AC PS9047;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT- LRR- and PYD-containing protein 5 (Water protein homolog).
 GN NALP5 OR WATER.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=21922687; PubMed=11925379;
 RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
 RT "A human homologue of mouse Mater, a maternal effect gene essential
 for early embryonic development.";
 RL Hum. Reprod. 17:903-911(2002).
 CC -!- TISSUE SPECIFICITY: Oocyte specific.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
 Mater.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AY054986; AAL15549.1; -;
 CC Genew; HGNC:21269; NALPS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR RNinh.
 DR InterPro; IPR003590; LRR RNinh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 3.
 DR Pfam; PF05729; NACHT; 1.
 DR SMART; SM00368; LRR_R1; 11.
 DR PROSITE; PSS0824; DAPIN; 1.
 DR PROSITE; PSS0837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 57 148 DAPIN.
 FT DOMAIN 280 602 NACHT.
 FT REPEAT 704 727 LRR 1.
 FT REPEAT 730 753 LRR 2.
 FT REPEAT 780 803 LRR 3.
 FT REPEAT 809 832 LRR 4.
 FT REPEAT 836 863 LRR 5.
 FT REPEAT 865 892 LRR 6.
 FT REPEAT 893 916 LRR 7.
 FT REPEAT 950 973 LRR 8.
 FT REPEAT 979 1002 LRR 9.
 FT REPEAT 1007 1034 LRR 10.
 FT REPEAT 1036 1059 LRR 11.
 FT REPEAT 1064 1092 LRR 12.
 FT REPEAT 1121 1142 LRR 13.
 FT NP_BIND 286 293 ATP (POTENTIAL).
 SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A71B28FA CRC64;
 Query Match 24.1%; Score 1320; DB 1; Length 1200;
 Best Local Similarity 32.7%; Pred. No. 1.6e-87;
 Matches 351; Conservative 167; Mismatches 461; Indels 96; Gaps 18;
 QY 17 LLEFAVELKKFK-LYLGTATLGECKIPWGSMEKAGPLEMAQLLIITHCPEANFLALS 75
 DB 69 LYLDEKEEFQTFKKKKSSSTTCISPFQFIENANVECLALLLHEYTGASLWATSIS 128
 QY 76 TFERINRDLWRGOREDLVR---DPQETRYDYVRKFKRL----- 112
 DB 129 IFENMMLTISEKA-RDDMKRHSPEPEATMTDQGPSKEKVPGISQAVQDSATAETKE 187
 QY 113 -----MEDRNARLGEVNLSHRYTLLLVKHSNPMQVQQQLLDTGRGHARTVGHQASP 166
 DB 188 QEISQAMEQEGATAA--TEEQEISQAMEQEGATAAETEEQ-----GHGGDTWDYKSH 238
 QY 167 IKETLTFEPEDEER-----PE-----PPRTVVMQGAAGIGKSLAHK 202
 DB 239 VMTKFAEEEDVRRSFENTAADPEMOTLAGAFDSDRMGPRPRTVVLHGKSGIGKSALARR 298

QY 203 VMLDWADGKLFQGRDYLYFYINCREMNOSETECSMODLIFSCWPEPSAPLOELIRVPERL 262
 DB 299 IVLVWAQGLVQGMESYVFFLPVREM-QRKKESSVTEFLSRWPDSPQAPVTEIMSRPERL 357
 QY 263 LFIIDGFDLPSFHDPOQPMCLCWEEKRPTTELLNSLRKLLPELSLIITRTALRK 322
 DB 358 LFIIDGFDL--GSVLNNDTKLCKDWAQKPPPTLRSLLRKLVPESFLIVTRVDGTEK 416
 QY 323 LHRLEHPRHVEILGFSAEAEKEYKYKFHNAEAGOVNRYVDNEPLFTMCFVPLVCWV 382
 DB 417 LKSEVVSRYLLVRGIGSGEQRHLLERIGISHQTKQGLRAINNRELLDQCVFVAVGSL 476
 QY 383 VCTCLQQQLGEGGLLRQTSRTTAVVLYLLSLMOPKGPAPR---LQPPNORGLCSLAA 439
 DB 477 ICVALQLQDVGESVAPNQTLTGHAFAFQHLTPRGVVRCLNLEERVVLKRCFMVAV 536
 QY 440 DGLMNQKILFEQDILRKHLGDEGVSFANMIFOKDINCERYYSFIHLSFQEFFAAMY 499
 DB 537 EGVNKRKSVFDGDDLMVQGLGESELRALPHNMILLPDSHCEEYVTFHLSLQDFCAALY 596
 QY 500 ILDEGEGGAGPDODVTRLLTEYAFSEFSE-----LALTSFLGLLNEETRSHE 549
 DB 597 VLE-----GLEIEPALCPLYVEK--TKRSMELKQAGFHIHSLWMKRFPLGNSVEDVRPLE 650
 QY 550 KSLCWKVSPIHAKMDLLQMTQSKAQSDGSTLOQGSLEFFSCLYEIOEEFIQALSHFOVI 609
 DB 651 VLLGCFVPLGVKQLLHWVSLGQPNATTPGDTLDAPHCLFETQDKFVRLALNSFQEV 710
 QY 610 VVSNIASKMEHVVSVFCLKRCRCSAQLVHLHYGATYSADGEDRARCAGANTLIVQLRPERT 669
 DB 711 WLP-INQNLDLIASSFCLQHCFYLRKIRVDVKGIFPRDESABACP-----VVPLWMRDK 763
 QY 670 VLLDAYSEHLAALCTNPNIELSYRNALSGRVKLLCQGLRHHPKCLONLRKCRIS 729
 DB 764 TLIEQWEDFCSMGLGTHPLHRLDGLSSILTERAMKTLCAKURHPTCKIQTLMFRNAQIT 823
 QY 730 SSACEDLSAALANKNLTRMDLSGNGVGPFGNMMLCEGLRHHPQCRLOMTQLRKCOLESGA 789
 DB 824 PGV-QHLMRWIVMANRNLRSNLGLGTHLKEEDVRMACEALKHPKCLLESRLDCCGLTHAC 882
 QY 790 COEMASVLGTNPHILVELDTGNALDELGLRLCQGLRHHPVCLRTLWLKILKLTAAACDE 849
 DB 883 YLKISQILTTSPSLKSLAGNKTQDQVTPLSDALRVSQCALQKLEDCGITTATGCOSS 942
 QY 850 LASTLSVNOQLRELDLSNELGDLVLLCEGLRHPTCKIQTLRILGICRLGSAACEGLSV 909
 DB 943 LASALVSNRSLTHLCLSNNSLNEGVLNLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002
 QY 910 VILQANHNLRELDLSFNDLGDWGLWLAEGLOHPACRLOKILWLDSCGLTAKACENLYFTLG 969
 DB 1003 ALMGNSWLTHLSLNMNVEDNGVKLLCEVMRPSCHLQDLVVKCHLTAACCELSLSCVIS 1062
 QY 970 INQTLTDLYTNALGDTGVRLLCKRLSHPGCKLRVLMFGMDLNKMTSHRLAAL 1024
 DB 1063 RSRHLKSLDLDNALGDGVAALCEGLKQ---KNSVLTRLGLKACGLTSDCCBAL 1114

RESULT 8

NAL2_HUMAN STANDARD; PRT; 1062 AA.
 ID Q9NKO2; Q9BNV5; Q9H6G6; Q9HAV9; Q9NWK3;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 2 (PYRIN-containing APAF1-like
 protein 2) (Nucleotide-binding site protein 1).
 GN NALP2 OR PYPAF2 OR NBS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]

CC	or send an email to license@isb-sib.ch .	
CC		
DR	EMBL; AF298547; AGI5253.1; ALT_INIT.	
DR	EMBL; AF310106; AAG30289.1; -	
DR	EMBL; AF464764; AAL69962.1; -	
DR	EMBL; AK000517; BAA91223.1; -	
DR	EMBL; AK025952; BAH15293.1; -	
DR	EMBL; AK000784; BAA91377.1; ALT_INIT.	
DR	EMBL; BC003592; AAH03592.1; -	
DR	EMBL; BC001039; AAH01039.1; -	
DR	HSP; PI0775; 2ENH.	
DR	Genew; HGNC:22948; NALP2.	
DR	InterPro; IPR007091; LRR_Rninh.	
DR	InterPro; IPR007111; NACHT_NTPase.	
DR	InterPro; IPR004020; PAAD_DAPIN_dom.	
DR	PFam; PF05729; NACHT; 1.	
DR	PROSITE; PSS0824; DAPIN; 1.	
DR	PROSITE; PSS0837; NACHT; 1.	
KW	Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;	
KW	Alternative splicing.	
FT	DOMAIN 1	DAPIN.
FT	DOMAIN 207	NACHT.
FT	REPEAT 467	LRR 1.
FT	REPEAT 622	LRR 2.
FT	REPEAT 754	LRR 3.
FT	REPEAT 810	LRR 4.
FT	REPEAT 839	LRR 5.
FT	REPEAT 867	LRR 6.
FT	REPEAT 924	LRR 7.
FT	REPEAT 981	LRR 8.
FT	REPEAT 1010	LRR 9.
FT	NP BIND 213	ATP (POTENTIAL).
FT	DOMAIN 518	POLY-GLU.
FT	VARSPPLIC 133	Missing (in isoform 2).
FT	CONFLICT 1	/FTID=VSP_005522.
FT	CONFLICT 35	M -> V (IN REF. 1).
FT	CONFLICT 304	L -> P (IN REF. 1).
FT	CONFLICT 364	I -> S (IN REF. 4; BAB15293).
FT	CONFLICT 980	R -> K (IN REF. 1).
FT	CONFLICT 1052	MISSING (IN REF. 1).
FT	CONFLICT 1062	A -> E (IN REF. 5).
SEQ	SEQUENCE 1062 AA; 120514 MW; 4DBB0F6E9C2BC8A7 CRC64;	
	Query Match 23.9%; Score 1309; DB 1; Length 1062;	
	Best Local Similarity 33.0%; Pred. No. 8.5e-87;	
	Matches 352; Conservative 176; Mismatches 430; Indels 108; Gaps 24;	
Qy	13 LSTVLELEAVELVKFKLYLGATELGE--GKIPWGSMEKAGPLEMAQLLTHFGPEAW 70	
Db	11 LQALLEQLSQDELKFKK-YLIITFSLAHELYQKIPHEKVDKADGQKQVLEILLTHCDSYWE 69	
Qy	71 RLALSTFEIRNRKDLWERGQRELDVRDPQFTYRDYVRKFKELMDRNLARGECVNLSHRY 130	
Db	70 MASLQVFEKHMRLDLSERKDE-----VREALKSF-----NKRKPLSLGITRKE 114	
Qy	131 TRLLLVKEHNSPMQVQOQLDLTGCHARTVGHQA---SPTK-----IETLEE----- 174	
Db	115 RPLDLDVDEMLERFKTEAQAFTETKGNVICLGKEVFKGKKPKDKNRCRYLTKTKFERMWS 174	
Qy	175 -PDEER-----PEP-PRTVVMQGAAGIGKSMIAHKWMLMDWADGK 211	
Db	175 WPGDSKEVQVMAERYKMLIPSPNPRVLPGFYSYTVLVGYPAGLGKTTLAQKMLMDWAEDN 234	
Qy	212 LFQGRFDYLVINCENMQSATECSMODLIFSCHWPESAPLOELIYRPERLLFIIDGFE 271	
Db	235 LIH-KFYAFYLSCRELSRLG-PCSFALVFRDWPQLQDDIPHTLAQKRLFIIDGFE 292	
Qy	272 LKPSFDHPQGPWCLCWEEKRPTLLNSLRKLLPELSLILITTRPTALEKHLLEHPR 331	
Db	293 LGAAPGALIEDICGDWEKKFVPVLLGSLNRMVLPKAALLVTTTPRALRDLRIAEPI 352	
Qy	332 HVEILGFSEARKEYFYKFNHAEQAQGVFNVYVRDNEPLFTFMCFVPLVCVWVCTCLOQOL 391	

Db 353 YIRVEGFEEDRRAYFLRHFGDEQAMRAFAELMRSNAALFOLGSAPAVCVIWTCTKLQM 412
 QY 322 EGGGLLRQTSRTTAVMYLLSLMQPKGAPLQPPNORGLCSLAADGLWKNOKILPEE 451
 Db 413 EKXGDFVPTCLRTGLFLRFLSRF---PQGAQLRGA--LRTLSLIAAQLMAQTSVLHR 467
 QY 452 QDLRKHGDLGDEVSAFLNNWIFQKINCERYYSFTHLSFQEPFAAMYYILDSGE---G 506
 Db 468 EDLERLGVQESDRLRFLDGLIURQDVSKGVSFTHLSFQFLTAFTYTLKEBEDRGG 527
 QY 507 GAGPQDDVTRLLTEYAFSERSFIATLSRFLGLNEETRSHEKSLCWKVSPIHKMDLQ 566
 Db 528 HTWDIGDVQKLISGVERLRNPDLIQAGYYSFGLANEKRAKELEATFGCRMSPIQKELL 587
 QY 567 W-IQSK-AQSDGSTLOQGSLEFFSCYIEIOBEFTQOALSHPQVIIVVNSIAKMEHWS 624
 Db 588 CDISCKGGHSTVTDLQ---ELLGCLYESQEBELKVEVMAQFKEI--SLHLNAVDPVPS 641
 QY 625 FCLKCRSAQVHLHY-----GATYSADGDRARCSAGAHVLLVQLRPERTVLLDAYS 676
 Db 642 FCVKHCRLNQLKMSLQVICKENLPENVTASDEAVERSQDDQHWL-----PFW 688
 QY 677 EHLAALCTNPNIELSYRNALSGRGVKKLQGLRHPNCKLQNLKRCRISSSACEDL 736
 Db 689 TDLCSIFGSKDMLGLAINDSFLSALVRILCEQIASDTCHLQVVFKNIS-PADAHRL 747
 QY 737 SAALANKVLTMDLSGNGV--FPGMMLLCEGLRHPQRLQMIQLRKQLESAGCQEMA 794
 Db 748 CLALRGHKVTYVTLTGNDQDDMFPA--LCEVLRHPECNRLRYLGLVSGSATTTQOWADLS 804
 QY 795 SVLTGNPHVAVELDTGNALEDLGLRLCOGLRHPVCRLRTLMKICRLTAACDELASTL 854
 Db 805 LALEVNSQSTCVNLSNELLDSGAKLLYTLRHPKCFLORLSLENCHLTAENCKDLAAVL 864
 QY 855 SYNQSLRELDLSINELGDLVLLLCBGRHPCTKLTQLRLGICRLGSAACBGLSVVLQAN 914
 Db 865 VVSRELTHLCLAKNPIGNTVKFLCBGLRYPECKLTQLVLMNCIDTSDGCCDLTKLQEK 924
 QY 915 HNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKACENLYTLGINOTL 974
 Db 925 SLLCLDLGLNHHGVKGMKFLCEALRKPLCNRLCMLWGCSPFPFCECDLCSALSCNQL 984
 QY 975 TDLYLTNNALGTVRLKCKLSHPGCKLRVLWL-----FGMDLNKM 1016
 Db 985 VTLDLQNPFLSSGVKMLFETLTCSSGTTLRLKLDIDFDELNKL 1030

RESULT 9
 ID NAL7 HUMAN STANDARD; PRT; 980 AA.
 AC Q8WY94.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 7 (PYRIN-containing APAFI-like protein 3).
 GN NALP7 OR PYPAF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.,
 RT "PYPAF7, a novel PYRIN-containing AFAFI-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing";
 RL J. Biol. Chem. 277:29874-29880 (2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22451042; PubMed=12563287;
 RX

RA Tschopp J., Martinon F., Burns K.;
 RT "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104 (2003).
 CC -|- SIMILARITY: Contains 1 DAPIN domain.
 CC -|- SIMILARITY: Contains 1 NACHT domain.
 CC -|- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF464765; AAL69963.1; -;
 DR EMBL; AY154462; AAO18158.1; -;
 DR GenBank; HGNC:22947; NALP7.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 93
 FT DAPIN.
 FT DOMAIN 172 491
 FT NACHT.
 FT REPEAT 614 638
 FT LRR 1.
 FT REPEAT 674 697
 FT LRR 2.
 FT REPEAT 760 784
 FT LRR 3.
 FT REPEAT 788 810
 FT LRR 4.
 FT REPEAT 817 840
 FT LRR 5.
 FT REPEAT 845 868
 FT LRR 6.
 FT REPEAT 874 897
 FT LRR 7.
 FT REPEAT 902 928
 FT LRR 8.
 FT REPEAT 933 957
 FT LRR 9.
 FT NP BIND 178 185
 FT ATP
 SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;

Query Match 21.3%; Score 1163; DB 1; Length 980;
 Best Local Similarity 33.8%; Pred. No. 2.9e-76;
 Matches 333; Conservative 147; Mismatches 403; Indels 102; Gaps 30;

QY 13 LSTYLELEAEVLEKFKLVLTATLGE--GKIPGSMKAGKAPLEMAQLLTHFGPEEAW 70
 Db 10 LQTLLEQNEDELKSKSL--WAFPLEDLQKTPWSEVEADCKLAEILVN--TSSEW 66
 QY 71 --RLALSTPERINRKDMERGQEDLVDPQRTYRDYVRRKFLMEDRNARLGEVNLSH 128
 Db 67 IRNATVNIILEENLTCLCKMAKAE--MMEDGQ-----VQIDNPGLGDAEEDSE 113
 QY 129 RYTRLLLVKEHS-----NPMQVQOQLL-----DTGRGHAR-TVCHQASPIKIEITLFP 175
 Db 114 -----LAKPGKEGWRNSMEKQSLVWKNTFWQGDIDNFHDVDTLENQ-----RFIFPLNP 163
 QY 176 DEERPEPRPTVYVQGAAGIKGKMLAHKVMLDWADGKLFQGRFDYLYINCREMNQSAEC 235
 Db 164 RTRPKLTPTVTVLHGPGVGTLLAKKMLDWTDCNL-SPTLYAFYLSCKELSRMG-PC 221
 QY 236 SMDLIIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGPW----CLCWEKKR 291
 Db 222 SFAELISKDWPELQDDIPISLAQAQRILFWVDGLDELKV----PPGALLQDIDGDEWKKK 277
 QY 292 PTELLNSLIRKLLPELSLLITTPRTALEKHLRLEHPRHVEILLGFSAAEKEVYKF 351
 Db 278 PVPVLGSLKLLKMLPRAALLVTTRALRDQLLAQQPIYVRVEGFLBEDRRAYFLRRF 337
 QY 352 HNAEQAGQVFNVRDNEPLFTMCFVPLVCVWVCTCLQQOLEGGLLRQTSRTTAYMYL 411
 Db 338 GDEDQAMRAFAELMRSNAALFOLGSAPAVCVIWTCTKLQWKEGDDVPVCLTFTGLFLRF 397
 QY 412 LLSLMQPKGAPRLQPPNORGLCSLAADGLWKNOKILFEEQDLRKHGLDGEDVSAFLNMN 471

Db 398 LCSRF---PQAQLRGA--LRTSLAAQGLWAQMSVPHREDLERLGVQESDLRLFLDGD 452
 Qy 472 IFKDIHCERYYSFTHLSFQFFFAAMYYIL--DEGE---GGAGPDQDDVTRLLTYAFESR 526
 Db 453 ILRQDRVSKGCSYFTHLSFQFLATFALFKEGEDRDGHAWDIDGVQKLLSGEERLKN 512
 Qy 527 SFALTSRFLPGLNEETRSHLEKSLCWKSPHIMDLLOWIQSKAQSDGSTLOOGLSLEF 586
 Db 513 PDLIQVGHFLPGLNAEKRAKELEATFGCRMSPDIKQELLQ--CKAHLHANKPLSVTDLKEV 571
 Qy 587 FSCLYEIQEERFIQALSHFQVIVVSNATSKMEHVMVSFCLKRCESAOVLHLYGA----- 641
 Db 572 LGCYSESQEEELAKVVVAPFKEISL-HUTINSEVMHCSFLKHCDLQKLSLOVAKGVFL 630
 Qy 642 ----TYSADGEDRARCAGATLWQLRPETVLLDAYSEHLAAALC-----TNPNIELS 693
 Db 631 ENYWDFFELDIEFE-RCT-----YLTIPNWARQDLRSRLWTDFCSLFSSNSNKLFE 681
 Qy 694 LYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALANKLITRMDLSG 753
 Db 682 VKQSFSLDSSVRILCDHVTSTCHLQKVEIKNV-TPDTAYRDFCLAFIGKKTLLTHTLAG 740
 Qy 754 NGVGFPGMML-LCEGLRHPQCRLOMIQLRKQLESAC-----QEMASVLGTNPHLVEL 806
 Db 741 HIWERTNMLMCDLRLRHKNLQVIRL-----GGHCATPEQWAEFFYVLKANQSLKHL 794
 Qy 807 DLTGNALEDGLRLCOGLRHPVCRRLTLWLKICRLTAAACDELASTLSVNQSLRELDLS 866
 Db 795 RLSANVLLDEGAMLLYKMTKPRKFLQMLSLNCRLEASCKDLAAVLVSVKLLTHLCIA 854
 Qy 867 LNEIGDLGVLLCEGLRHPTCKLQTLRIGICRLGSAACEGLSVVLQAHNLRBELDLSND 926
 Db 855 KNPIGDTGVKFLCEGLSPYDCKLQTLVLQOCSITKLGCRYLSEALQEAACSLTNLDLSINQ 914
 Qy 927 LGDMGLWLLAELGLOHPACRLQKWL 951
 Db 915 IAR-GLWILCOALENPNCLKHLRL 938

RESULT 10

NAL5 MOUSE
 ID NAL5 MOUSE STANDARD; PRT; 1111 AA.
 AC Q9RIM5; Q9ULR2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 5 (Maternal antigen that
 DE embryos require) (Mater protein) (Ooplasm-specific protein 1) (Op1).
 GN NALP5 OR MATER.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH Swiss; TISSUE=Ovary;
 RX MEDLINE=99360614; PubMed=10433232;
 RA Tong Z.-B., Nelson L.M.;
 RT "A mouse gene encoding an oocyte antigen associated with autoimmune
 RT premature ovarian failure.";
 RL Endocrinology 140:3720-3726(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS ASP-7; LEU-42; LYS-87; PRO-256;
 RP GLN-977 AND 1003-ASN-ASN-1004.
 RC STRAIN=129/SV;
 RX MEDLINE=20222768; PubMed=10754103;
 RA Tong Z.-B., Nelson L.M., Dean J.;
 RT "Mater encodes a maternal protein in mice with a leucine-rich repeat
 RT domain homologous to porcine ribonuclease inhibitor.";
 RL Mamm. Genome 11:281-287(2000).
 RN [3]
 RP FUNCTION.

EX MEDLINE=20517328; PubMed=11062459;
 RA Tong Z.-B., Gold L., Pfeifer K.E., Dorward H., Lee E., Bondy C.A.,
 RA Dean J., Nelson L.M.;
 RT "Mater, a maternal effect gene required for early embryonic
 RT development in mice.";
 RL Nat. Genet. 26:267-268(2000).
 CC -!- FUNCTION: Necessary for embryonic development beyond the 2-cell
 CC stage.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Oocyte specific.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF074018; AAD51762.1; -.
 DR EMBL; AF143573; AAF64393.1; -.
 DR EMBL; AF143559; AAF64393.1; JOINED.
 DR EMBL; AF143560; AAF64393.1; JOINED.
 DR EMBL; AF143581; AAF64393.1; JOINED.
 DR EMBL; AF143562; AAF64393.1; JOINED.
 DR EMBL; AF143563; AAF64393.1; JOINED.
 DR EMBL; AF143564; AAF64393.1; JOINED.
 DR EMBL; AF143565; AAF64393.1; JOINED.
 DR EMBL; AF143566; AAF64393.1; JOINED.
 DR EMBL; AF143567; AAF64393.1; JOINED.
 DR EMBL; AF143568; AAF64393.1; JOINED.
 DR EMBL; AF143569; AAF64393.1; JOINED.
 DR EMBL; AF143570; AAF64393.1; JOINED.
 DR EMBL; AF143571; AAF64393.1; JOINED.
 DR EMBL; AF143572; AAF64393.1; JOINED.
 DR PIR; A59000; A59000.
 DR HSP; P10775; 2BNH.
 DR MGD; MGI:1345193; Mater.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR GO; GO:0009887; P:organogenesis; IMP.
 DR InterPro; IPR001611; LRR_RNinh.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTFase.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat; Polymorphism.
 FT DOMAIN 1 128
 FT REPEAT 1 23
 FT REPEAT 24 49
 FT REPEAT 50 75
 FT REPEAT 76 102
 FT REPEAT 103 128
 FT REPEAT 129 513
 FT REPEAT 514 664
 FT REPEAT 665 714
 FT REPEAT 715 762
 FT REPEAT 763 799
 FT REPEAT 800 832
 FT REPEAT 833 856
 FT REPEAT 857 884
 FT REPEAT 885 913
 FT REPEAT 914 941
 FT REPEAT 942 974
 FT REPEAT 975 1003
 FT REPEAT 1004 1027
 FT REPEAT 1028 1059
 FT NP BIND 197 204
 FT VARIANT 7 7
 FT VARIANT 42 42
 FT VARIANT 87 87
 E -> D (IN STRAIN 129/SV).
 G -> L (IN STRAIN 129/SV).
 R -> K (IN STRAIN 129/SV).


```

FT DOMAIN 196 513 NACHT.
FT DOMAIN 604 614 POLY-GLU.
FT REPEAT 462 487 LRR 1.
FT REPEAT 727 747 LRR 2.
FT REPEAT 755 778 LRR 3.
FT REPEAT 811 834 LRR 4.
FT REPEAT 845 868 LRR 5.
FT NP BIND 202 209 ATP (POTENTIAL).
SQ SEQUENCE 892 AA; 98733 MW; 4AABD1FC766DE9D CRC64;

Query Match
Best Local Similarity 21.2%; Score 1159; DB 1; Length 892;
Matches 301; Conservative 155; Mismatches 340; Indels 128; Gaps 20;

QY 17 LEELEAVLKKFKLYLGATGELGKIEPWGMEKAGPLEMAQLLTHFGPEANRLALST 76
D 26 LEELEAVEQLKFRHKLKRDVGPGR-SIPWGRLEADAVDLAEQLAQFYGPPEALEVARKT 84
QY 77 FERINRKLWEGQREDL-----VRDQETRDYVRKFRMLMEDNRARLGCY 124
D 85 LKRAADARVAQAQLQRRLQRLGLSGTLLSVSEYKKYREHVLQHAHVKNAR---SV 141
QY 125 NLSHRYTLVLLVKEHNSPMQVQQLDTRGHARTVGHQASPIKLTETLFEDEERPEPR 184
D 142 KITKFTKLLIAPESAAPPEALGPAEPEPGAR-----RSDTHFNRLFRDEE-GRPL 196
QY 195 TVVQGAAGIGKSLAHKVMLDWADKLFQGRFDYLYFVINCENMQSATECSMDLIFSC 244
D 197 TVVLQGPAGIGKTAARKILYDWAAGKLYQGQVDFAFWPCGELLERPGTRSLADLILDQ 256
QY 245 WPESAPLOELIRVPERLLFIIDGDELKPSHDPQGWCLWEEKRETELLINSLRKK 304
D 257 CPDRCAPQMLAQQRLLFLDGDDEL-PALGGPEAAPCTDPPFEAASGARVLGLLSKA 315
QY 305 LLPESLLITRPTALEKHLRLEHPRVEILGFSEAEKRYFYFHNAAQGVFNKV 364
D 316 LLPALLLVTTAAAPGRGLGELCSFQCAEVGFSDDKDKKYFYFFDERAEARVRF 375
QY 365 RDNEPLFTMCFVPLVWVCTCLOOLEGGGLLRTSTRTTAVMYLLSLMQPKGA-- 422
D 376 KENETLFCALCFVFCVWVCTVLRQLELGRDLRSRTKTTTSLVLLFTITVSSAPVADG 435
QY 423 PRLQPPNQRGLCSLAADGLWNOKILFEOQLRKHGLGDEV-SAFINMNFQKDINCER 481
D 436 PRLQ--GDLNRNCLAREGLVGRRAQFAEKELEQLRGSKVQTLFLSKKELPGVLETEV 493
QY 482 YVSFTHLSFQEFFAAMYILDEG-----EGAGPDQDVTRLLTETAFSEFSFALTSTRF 535
D 494 TYQFIDQSFFQFLAALSYLEDDGVPRTAAGVG-----TLRGDAQPHSHLVITTRF 546
QY 536 LFGLLNEETRSHLEKSLCWKYSPIHMDLLOWIQSKAQ-----SDG 576
D 547 LFGLLSAERMRDIERHFGWMSERVQKQALRVWVGQGGCPGVAPEVTEGAKGLEDTPEP 606
QY 577 STLQGS-----LRFPSCLYETQEEFTQOALSHFQIVVSNIA-SKMEHWVSSFCLKRC 630
D 607 EEEEBEENPPELLELYCYETQEDAFVROALQRPPELALQVRRCMDVAVLSTVCRCC 666
QY 631 RSAQVHLVYATYSADGEDRACSAHAHTLLVQLRPERTVLLDAYSEHLAAALCTNPNI 690
D 667 PAGQALRLTSCRLVAQEKKKK-----SLGKRLQASL----- 698
QY 691 ELISLYRNALGSRG-----VKLLCQGLRHPNCKLQNLKRCRISSACEDLSAALI 741
D 699 -----GGGSSSQGTTKQLPASLLHPLFQAMTDPLCHLSLTLSSHCKLPDAVCRDLSEALR 753
QY 742 ANKNLTRMDLSCNGVGFPCPMMLLCEGLRHPQCLQMIQLRKQCSGACQENASVLGTNP 801
D 754 AAPALTEGLLHNRUSEAGLRMLSEGLAMPQCRQVTRVQLPDPQRG-LQYLVGMLRSP 812
QY 802 HIVELDLTGNALDELGLRLCCGLRHPVCRRLTLMKICRLTAAACDELAFLSTLVNQSLR 861
D 813 ALTTLDLSCQLPAPMVTYVLCVLAHQGGGLQTL-----SLASVELSEQSLQ 859

```

```

QY 862 ELDLSLNELGDLGVLLICEGLRHP 885
D 860 ELQAVKRAKPDVLI-----THP 876

RESULT 12
NAL6_MOUSE STANDARD; PRT; 843 AA.
AC Q91WS2; Q8K0L4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE PYRIN-containing APAF1-like protein 5-like.
OS NALP6 OR PYPAF5.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP CONCEPTUAL TRANSLATION OF 1-287.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RP SEQUENCE OF 288-843 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RX MEDLINE=22520874; PubMed=12633874;
RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT "Identification of mammalian orthologs associated PYPAF5 with distinct
RT functional roles.";
RL FEBS Lett. 538:173-177 (2003).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B (By similarity).
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC sequences, in analogy to ortholog sequences.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC013519; AAH13519.1; -

```


RINI_HUMAN
ID RINI_HUMAN STANDARD; PRT; 460 AA.
AC P13489;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental ribonuclease inhibitor (Ribonuclease/angiogenin inhibitor)
DE (RAI) (RNase inhibitor) (RI).
GN RNH OR PRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118269; PubMed=3219362;
RA Lee F.S., Fox E.A., Zhou H.-M., Strydom D.J., Vallee B.L.;
RT "Primary structure of human placental ribonuclease inhibitor.";
RL Biochemistry 27:8545-8553(1988).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 443-461.
RX MEDLINE=89210799; PubMed=3243277;
RA Schneider R., Schneider-Scherzer E., Thurnher M., Auer B.,
RA Schweiger M.;
RT "The primary structure of human ribonuclease/angiogenin inhibitor
(RAI) discloses a novel highly diversified protein superfamily with a
common repetitive module.";
RL EMBO J. 7:4151-4156(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE-Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE-Brain, Kidney, Lymph, and Ovary;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussidi T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF COMPLEX WITH ANGIOGENIN.
RX MEDLINE=97459904; PubMed=9311977;
RA Papageorgiou A.C., Shapiro R., Acharya K.R.;
RT "Molecular recognition of human angiogenin by placental ribonuclease
inhibitor -- an X-ray crystallographic study at 2.0-A resolution.";
RL EMBO J. 16:5162-5177(1997).
CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
function in the modulation of cellular activities.
CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC DR EMBL; M22414; AAA59130.1; -;
CC DR EMBL; X13973; CAA32151.1; -;
CC DR EMBL; M36717; AAM60249.1; -;
CC DR EMBL; AL161967; CAB82310.1; -;
CC DR EMBL; BC003075; AAH03075.1; -;
CC DR EMBL; BC011500; AAH11500.1; -;
CC DR EMBL; BC014629; AAH14629.1; -;
CC DR EMBL; BC047730; AAH47730.1; -;
CC DR PIR; A31858; A31858.
CC DR PDB; 1A4Y; 14-OCT-98.
CC DR Genew; HGNC:10074; RNH.
CC MIM; 173320; -;
CC GO; GO:0008428; P:ribonuclease inhibitor activity; TAS.
CC GO; GO:0006401; P:RNA catabolism; TAS.
CC DR InterPro; IPR001611; LRR_RNinh.
CC DR InterPro; IPR007091; LRR_RNinh.
CC DR Pfam; PF00560; LRR; 3.
CC DR SMART; SM00368; LRR_RI; 1.
KW Repeat; Leucine-rich repeat; 3D-structure; Polymorphism.
FT INIT_MET 0 0
FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
FT REPEAT 19 47 LRR A1.
FT REPEAT 48 75 LRR B1.
FT REPEAT 76 104 LRR A2.
FT REPEAT 105 132 LRR B2.
FT REPEAT 133 161 LRR A3.
FT REPEAT 162 189 LRR B3.
FT REPEAT 190 218 LRR A4.
FT REPEAT 219 246 LRR B4.
FT REPEAT 247 275 LRR A5.
FT REPEAT 276 303 LRR B5.
FT REPEAT 304 332 LRR A6.
FT REPEAT 333 360 LRR B6.
FT REPEAT 361 389 LRR A7.
FT REPEAT 390 417 LRR B7.
FT REPEAT 418 446 LRR A8.
FT VARIANT 169 169 P -> L (in dbSNP:17585).
FT CONFLICT 422 423 /FTId=VAR_014726.
FT STRAND 2 10 RQ -> SE (IN REF. 2).
FT HELIX 16 22
FT TURN 23 28
FT STRAND 31 35
FT TURN 41 43
FT HELIX 44 52
FT TURN 53 53
FT STRAND 59 61
FT TURN 63 64
FT HELIX 68 76
FT TURN 77 79
FT TURN 82 83
FT STRAND 88 90
FT TURN 92 93
FT HELIX 98 100
FT TURN 101 110
FT TURN 112 113
FT STRAND 116 118
FT HELIX 125 136
FT TURN 137 137
FT TURN 139 140
FT STRAND 145 147
FT TURN 149 150
FT HELIX 155 157
FT TURN 158 167
FT TURN 169 170
FT STRAND 173 175
FT HELIX 182 195

Search completed: July 30, 2004, 13:48:14
Job time : 22 secs

FT	STRAND	202	204		
FT	TURN	206	207		
FT	TURN	213	213		
FT	HELIX	214	224		
FT	TURN	226	227		
FT	STRAND	230	232		
FT	HELIX	239	250		
FT	TURN	251	251		
FT	TURN	253	254		
FT	STRAND	259	261		
FT	TURN	263	264		
FT	HELIX	269	279		
FT	TURN	280	281		
FT	TURN	283	284		
FT	STRAND	287	289		
FT	TURN	291	292		
FT	HELIX	296	307		
FT	TURN	308	308		
FT	TURN	310	311		
FT	STRAND	316	318		
FT	TURN	320	321		
FT	STRAND	325	325		
FT	HELIX	326	328		
FT	HELIX	329	338		
FT	STRAND	344	346		
FT	STRAND	351	351		
FT	HELIX	353	364		
FT	TURN	365	365		
FT	TURN	367	368		
FT	STRAND	373	375		
FT	TURN	377	378		
FT	TURN	383	385		
FT	HELIX	386	395		
FT	STRAND	401	403		
FT	TURN	405	406		
FT	HELIX	411	422		
FT	TURN	424	425		
FT	STRAND	430	432		
FT	TURN	434	435		
FT	HELIX	440	452		
FT	TURN	454	455		
FT	STRAND	457	460		
SQ	SEQUENCE	460 AA; 49842 MW; C3D6668E2F2BF86F CRC64;			
Query Match 13.2%; Score 724; DB 1; Length 460;					
Best Local Similarity 42.5%; Pred. No. 6.6e-45;					
Matches 150; Conservative 64; Mismatches 139; Indels 0; Gaps 0;					
QY	671	LLDAYSEHLAAALCTNPNIIELSLYRNALGSRGVKLLCOGLRHPNCKLQNLRLKRCRISS	730		
DB	39	LTEARCKDISSALRYNPALAEINLRSNELGDVGVCVLOGLQTPSCIKIQLSLQNCCLAG	98		
QY	731	SACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQIRKQLES	790		
DB	99	AGCGVLSSTLRPTLTIQELHLSDNLLIGDAGLQCLCEGLLDPOCRLEKQLQLEYCSLSAASC	158		
QY	791	QRMASVILGNPHLVELDTGNALDELGLRLCOGLRHPVCRRLTLWLKICRLTAAACDEL	850		
DB	159	EPLASVIRAKPFKELTVSNNDINEAGVRVLOGLKDSQCQLEALKLES	218		
QY	851	ASTLSVNQSLRELDLSINELGDLGVLLCEGLRHPHTCKLQTLRLGICRLGSAACEGLSV	910		
DB	219	CGIVASKASIRELAGSNKLGVDVGMALCPGLLHPSSRLRLTWIWECCGITAKCGDLCRV	278		
QY	911	LQAHNLRLDLSFNDLGPWGLWLLAEGLOHPACRLQKMLDSCGLTAKACENLYETIGI	970		
DB	279	LRAKESLSELGAGNELGDEGARLLCETLLEPCQLES	338		
QY	971	NOTLTDLYLTNNALGDTGVRLLCRLSHPGCKLRVLWLFGMIDLNKMTSHRLAA	1023		
DB	339	NRFLLELQISNNELEDAGVRELCOGLGQPGSVLRVLWLADCDVSDSCSSLAA	391		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:46:06 ; Search time 47 Seconds
(without alignments)
6948.114 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 5472
Sequence: 1 MLRTAGDGLCRSLTYLEEL.....MTHSRALALRVTKPYLDIGC 1035

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4458.5	81.5	892	4 Q8NEU4	Q8neu4 homo sapien
2	1799	32.9	1093	4 Q8W24	Q8w24 homo sapien
3	1505.5	27.5	986	4 Q8W27	Q8w27 homo sapien
4	1450.5	26.5	994	4 Q8W87	Q8w87 homo sapien
5	1401	25.6	1043	4 Q8W25	Q8w25 homo sapien
6	1377	25.2	1375	4 Q8W25	Q8w25 homo sapien
7	1323	24.2	1200	4 Q8W29	Q8w29 homo sapien
8	1314.5	24.0	287	4 Q8Y26	Q8y26 homo sapien
9	1259	23.0	982	11 Q8U40	Q8u40 mus musculu
10	1216.5	22.2	1029	4 Q8W28	Q8w28 homo sapien
11	1168.5	21.4	1111	11 Q7TPU9	Q7tpu9 mus musculu
12	1062	19.4	713	6 Q9SLZ7	Q9slz7 macaca fasc
13	998.5	18.2	863	11 Q8C6J9	Q8c6j9 mus musculu
14	976	17.8	846	4 Q81XT0	Q81xt0 homo sapien
15	958	17.5	748	11 Q9MMW0	Q9mmw0 mus musculu
16	956	17.5	655	4 Q8W26	Q8w26 homo sapien

17	862.5	15.8	657	11 Q9BPG7	Q9bpg7 mus musculu
18	822	15.0	673	11 Q8CCN1	Q8ccn1 mus musculu
19	724	13.2	447	4 Q96FD7	Q96fd7 homo sapien
20	720	13.2	461	6 Q8H2P9	Q8h2p9 pan troglod
21	719	13.1	461	4 Q9BQ80	Q9bq80 homo sapien
22	717	13.1	456	11 Q91V17	Q91v17 mus musculu
23	716	13.1	456	11 Q924P4	Q924p4 mus musculu
24	711	13.0	461	4 Q81ZK8	Q81zk8 homo sapien
25	708	12.9	825	11 Q8C6M5	Q8c6m5 mus musculu
26	659	12.0	373	11 Q80YN6	Q80yn6 rattus norv
27	519	9.5	1013	11 Q80SZ8	Q80sz8 mus musculu
28	519	9.5	1013	11 Q80SY9	Q80sy9 mus musculu
29	519	9.5	1013	11 Q80SY8	Q80sy8 mus musculu
30	513	9.4	1013	11 Q80UH1	Q80uh1 mus musculu
31	513	9.4	1013	11 Q80SS7	Q80ss7 mus musculu
32	512	9.4	1013	11 Q80UH6	Q80uh6 mus musculu
33	512	9.4	1013	11 Q80UH5	Q80uh5 mus musculu
34	512	9.4	1013	11 Q80UH4	Q80uh4 mus musculu
35	512	9.4	1013	11 Q80SV7	Q80sv7 mus musculu
36	512	9.4	1013	11 Q80SV8	Q80sv8 mus musculu
37	511	9.3	1013	11 Q80UH7	Q80uh7 mus musculu
38	510	9.3	1013	11 Q80UH2	Q80uh2 mus spretus
39	507	9.3	1013	11 Q80UH3	Q80uh3 mus musculu
40	493.5	9.0	1866	4 Q86WI3	Q86wi3 homo sapien
41	464	8.5	269	11 Q9D458	Q9d458 mus musculu
42	462	8.4	390	11 Q8BWZ1	Q8bwz1 mus musculu
43	452	8.3	778	4 Q8NF48	Q8nf48 homo sapien
44	447.5	8.2	975	11 Q80W30	Q80w30 mus musculu
45	444	8.1	509	4 Q9H5Z8	Q9h5z8 homo sapien

ALIGNMENTS

RESULT 1

Q8NEU4	PRELIMINARY;	PRT;	892 AA.
ID Q8NEU4			
AC Q8NEU4;			
DT 01-OCT-2002 (Tremblrel. 22, Created)			
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE Monarch-1 splice form IV.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheraa; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Williams K.L., Linhoff M.W., Ting J.P.Y.;			
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY116207; AAM/5145.1; --			
DR InterPro; IPR001611; LRR.			
DR InterPro; IPR007091; LRR RNinh.			
DR InterPro; IPR007111; NACHT NTPase.			
DR InterPro; IPR004020; PAAD_DAPIN_dom.			
DR Pfam; PF00560; LRR; 1.			
DR Pfam; PF05729; NACHT; 1.			
DR Pfam; PF02758; PAAD_DAPIN; 1.			
DR PROSITE; PS50824; DAPIN; 1.			
DR PROSITE; PS50837; NACHT; 1.			
SQ SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;			

Query Match 81.5%; Score 4458.5; DB 4; Length 892;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 197; Gaps 2;

QY	1	MLRTAGDGLCRSLTYLEELKPKLYGTATLGEKIPWGSMEKAGLEMAQILL	60
DB	1	MLRTAGDGLCRSLTYLEELKPKLYGTATLGEKIPWGSMEKAGLEMAQILL	60
QY	61	ITHFGPEEAWRLALSTFERINRKDLWGERQEDLVR-----	96
DB	61	ITHFGPEEAWRLALSTFERINRKDLWGERQEDLVRDTPGGPSSLGNGSTCLLEVSLVT	120

QY 97 ---DPQETDYVRRKRLMEDRNARLGEVNLSHRYTLLLVKHSNPMQVQQQLDNG 153
Db 121 PKDQETDYVRRKRLMEDRNARLGEVNLSHRYTLLLVKHSNPMQVQQQLDNG 180
QY 154 RGHARTVGHQAQSPKIETLFEDEERPEPRPVVVMQGAAGIKSMLAHKVLMDWADGKLF 213
Db 181 RGHARTVGHQAQSPKIETLFEDEERPEPRPVVVMQGAAGIKSMLAHKVLMDWADGKLF 240
QY 214 QRFYDLFVINCENMQSATECMODLIPSCWPESAPLOELIRVPERLLFIIDGDEUK 273
Db 241 QRFYDLFVINCENMQSATECMODLIPSCWPESAPLOELIRVPERLLFIIDGDEUK 300
QY 274 PSFHPDQPGWCLCWEKRETELLNSLRKLLPELSLITRPTALEKHLRLEHPRIV 333
Db 301 PSFHPDQPGWCLCWEKRETELLNSLRKLLPELSLITRPTALEKHLRLEHPRIV 360
QY 334 ELIGFSEAEKKEYFYKIFHNAEQAGQVFNVRDNEPFTMCFVPLVWVCTCLOQQLRG 393
Db 361 ELIGFSEAEKKEYFYKIFHNAEQAGQVFNVRDNEPFTMCFVPLVWVCTCLOQQLRG 420
QY 394 GGLLRQTSRTTAVVWMLYLLSMQPKPGAPRLQPPNQRLGSLAADGLWNOKILFEED 453
Db 421 GGLLRQTSRTTAVVWMLYLLSMQPKPGAPRLQPPNQRLGSLAADGLWNOKILFEED 480
QY 454 LRKHGLDGEDVSAFLNMNIFQDINCERYYSFIHLSFQEFFAAMYVILDEGGGAGPDOD 513
Db 481 LRKHGLDGEDVSAFLNMNIFQDINCERYYSFIHLSFQEFFAAMYVILDEGGGAGPDOD 540
QY 514 VTRLLTEYAFSRSFALTSRFLGLNBEETSHLEKSLCWKVSPIHKMDLLQWISKAQ 573
Db 541 VTRLLTEYAFSRSFALTSRFLGLNBEETSHLEKSLCWKVSPIHKMDLLQWISKAQ 600
QY 574 SDGSTLOQSLFFSCLYHIOEBEFTQALSHQVIVVSNIAKSHVMVSSCLKRCRSA 633
Db 601 SDGSTLOQSLFFSCLYHIOEBEFTQALSHQVIVVSNIAKSHVMVSSCLKRCRSA 660
QY 634 QVHLHYGATYSADGEDRARCAGANTLLVQLRPTVLLDDAYSEHLAALCTNPNLIUS 693
Db 661 QVHLHYGATYSADGEDRARCAGANTLLVQLRPTVLLDDAYSEHLAALCTNPNLIUS 720
QY 694 LYRNALSGRGVLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 753
Db 721 LYRNALSGRGVLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 780
QY 754 NGVGFPGWMLLCEGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 813
Db 781 NGVGFPGWMLLCEGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 840
QY 814 EDGLRLLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 873
Db 841 EDGLRLLLCOGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 883
QY 874 GVLLLCGLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSG 933
Db 864 --- 863
QY 934 LLAEGLOHPACHLQKLDLSDGLTAKACENLYFTLGINOTLDTLYTNALGDTGVRLIC 993
Db 864 --- 863
QY 994 KRLSHPGCKRLVLMFGMDLNKWTSHRSLAALRVTKPYLDIGC 1035
Db 864 -----WLFMGMDLNKWTSHRSLAALRVTKPYLDIGC 892

RESULT 2

Q86W24
ID Q86W24 PRELIMINARY; PRT; 1093 AA.
AC Q86W24;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE NALP14.
GN NALP14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Teschopp J., Martinon F., Burns K.;
RT "NALPS: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104 (2003).
DR EMBL; AY154469; AAO18165.1;
DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
DR GO; GO:0005488; P:binding; IEA.
DR GO; GO:0008610; P:transport; IEA.
DR InterPro; IPR002194; Chapteronin_TCP-1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_R1; 12.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
DR PROSITE; PS00837; NACHT; 1.
DR PROSITE; PS00995; TCP1.3; 1.
SQ SEQUENCE 1093 AA; 124732 MW; 124EEACE22A11D6F CRC64;
Query Match 32.9%; Score 1799; DB 4; Length 1093;
Best Local Similarity 37.4%; Pred. No. 4e-147;
Matches 413; Conservative 177; Mismatches 388; Indels 126; Gaps 15;
QY 13 LSTYLEEAEVLEAEKFKLYLTATLGEKGIPIWGSMEKAGPIEMAQLLITHFGEAWRL 72
Db 15 LLLYLEELNKEELNTFLKLTETMEPEHGLTPWNEVKARREDLANMKYIPGEKAWSV 74
QY 73 ALSTERINRKLWRRGORE-----DLVRDPOET-----YRDYVRRKF 110
Db 75 SLKIFGKMLKDLCEAKEEINWSAQITIGPDDAKAGETQEDQEAVLGDGTEYRNRKIF 134
QY 111 RLMEORNLGECVNLSHRYTLLLVKHSNPMQVQQQLDNGRGHARTVGHQAQSPKIET 170
Db 135 CITWOKSLAGAPEDPHHG-----IAEKDKL-----LE 163
QY 171 TLFEDEERPEPRPVVVMQGAAGIKSMLAHKVLMDWADGKLFQGRFDYLFYINCREMNQ 230
Db 164 HLFVDVVKTAQPIQVVLQGAAGVKTTLVRKAMLDWAGSLYQGRFKVYVYLNREINQ 223
QY 231 SATSCMODLIPSCWPESAPLOELIRVPERLLFIIDGDEUKPSFHPDQPGWCLCWEK 290
Db 224 -LKERSFAQLISKOWPSTEGPTEEMYPSSLLFIIDSFELNFAFEPEFALCEDWTOE 282
QY 291 RPTLELLNSLRKLLPELSLITRPTALEKHLRLEHPRIVLIPSEAEKKEYFYKY 350
Db 283 HVPFSLMSLLKRWMLPEASLVTTLTTSKRLKQLKNHHYVELLGMSEDAEEIYOF 342
QY 351 FHNAEQAGQVFNVRDNEPFTMCFVPLVWVCTCLOQQLGGGLLRQTSRTTAVVWL 410
Db 343 FEDKEWAMKVFSSLSKSNEMLFSCMQVPLVCWAACTCLAQMEKGGDVTLCQTTTALFTC 402
QY 411 YLLSLMQP-KPGAPRLQPPNQ---RGLCSLAADGLWNOKILFEEDLKHGLDGEDVSA 466
Db 403 YISSLFTPDGGSPSL---PNOAQLRRLCQAAKGIWTVYVYRENLRRLRGLTOSDVSS 459
QY 467 FLNNMIFQDINCERYYSFIHLSFQEFFAAMYVIL-----DEGEGAGPDQDVTRLITEYA 522
Db 460 FMDSNIIQDAEYENCYVFTLHVGEFFAAMYVLMKGSWEAGNPSQCPPEDLKSLLSQSTS 519
QY 523 PSERSFLALTSRFLGLNBEETSHLEKSLCWKVSPIHKMDLLQWISKAQSDGSLTQOG 582

Db 520 YKD-PHLMQKCFGLNEDRVKQLERTFNCKMSLKIKSKLLQCMVEVLGNSDYSFSLQ 578
QY 583 SLFFSCLYEIQEEFTQALSHFOVIVVNSIAKMEHWSFCLKRCRAQVHLHYGA- 641
Db 579 FLBFLHCLYEQDQAFISQAWRCFPKVAI-NICEKIHLLVSSFCCKRCCLRTIRLSVTV 637
QY 642 -----TYSAGEDRARC-----SAGAHLLVQLRPTVLLDAYSEHI----- 679
Db 638 VPEKKILKTSPTWGDTRTHCWQDLCSVLHT-----NEHLRELDLHNSLDKSAWN 691
QY 680 -----AAALCTPNLIELSLYRNALSGRVKLLCQ 709
Db 692 ILHHLRHPNCKLQKLLKFTFPDGCODISTSLHNKNLMLDLKGSIDGNGVKSICE 751
QY 710 GLRHPNCKLQNLRLKRCRCHISSACEDLSAALIANKNLTRMDLSNGVGFPGMMLLCEGLR 769
Db 752 ALKHEPECKLQTLRLSCNLTVPCCNLISNALIRSQSLIFLNLSTNNLLDDGVQLLCEALR 811
QY 770 HPQRLQMIQLRKQLESCAGOCMAVSLGTNPHLVLDLTGNALDELGLRLLCQGLRHPV 829
Db 812 HPKCYLERLSLESCGLTAGCEYLSLALISNKRLLTHCLADNLVLDGGVGLMSDALQHAQ 871
QY 830 CLRRLTLWIKICRLTAACDELAETSLVNSQSLRELDLSNELDGLVILLCEGLRHPTCKL 889
Db 872 CTLKSLVLRCHFTSLSSEYLSLTHLNSKSLTHLDLGSNWLQDNGVYKLLCDVFRHPSCNL 931
QY 890 QTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKL 949
Db 932 QDLMLGCVLTWACCLDLASVILNPNRSLRSLDGNLQDDGVKILCDALRYPNCNIQRL 991
QY 950 WLDSCGLTAKACENLYFTLGINQTLTDLTYTNALDGTGVRLLCKRLSHPGCKLRVLMJF 1009
Db 992 GLEYGCLSLCCQDLSSALICNKRLLKNLTQNTLGYEGIVKLYKLSPKCKLQVLGLC 1051
QY 1010 GMDLNKMTSHRLAALRVKPYLDI 1033
Db 1052 KEAFDEEAQKLEAVGVSNPHLII 1075

PRELIMINARY; PRT; 986 AA.

Q86W27 ID Q86W27 PRELIMINARY; PRT; 986 AA.
AC Q86W27; (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NALP9.
GN NALP9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=2451042; PubMed=12563287;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
DR EMBL; AY154464; A018160.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR007091; LRR RNinh.
DR InterPro; IPR003590; LRR RNinh sub.
DR InterPro; IPR007111; NACHT NTPase.
DR InterPro; IPR004020; PAAD DAPIN dom.
DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF02758; PAAD DAPIN; 1.
DR SMART; SM00368; LRR_R1; 10.

DR PROSITE; PS00615; C TYPE LECTIN_1; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
SQ SEQUENCE 986 AA; 112592 MW; B4248B46F70413B2 CRC64;
Query Match 27.5%; Score 1505.5; DB 4; Length 986;
Best Local Similarity 34.2%; Pred. No. 1.2e-121;
Matches 344; Conservative 193; Mismatches 393; Indels 85; Gaps 18;
QY 16 YLEELEAVELEKFKLYLGTATLGEKGK-IPWGSMEKAGPLEMAQOLLTHFCPEEAWRLAL 74
Db 14 YLKLREKEEFKFKELLKQPLEKPELPWPAAELKASKEDVAKLLDKHYGPKQAWEVTL 73
QY 75 STFERINRKLWEGQREDLVDPQETVRDVRKFKRLMEDNRNARLGCNLSLHRYTRL 134
Db 74 NLFTQINRKDLWTKAQEE--NRKNLNPRKMKETFLQINEKET----CLHVEPHFYKET 127
QY 135 LVKHSNPMQVOQLLDTGRGHARTVGHQASPIKETLFEPEDEPEPRPTVVMQAGAI 194
Db 128 MNFY-----KELNDAYTAAARR-----HTVLEGGDGI 156
QY 195 GKSMIAHKVMDLWADGKLFQGRDYLVFVINCENMNSQATSCMODLIFSCWPEPAPQOE 254
Db 157 KKTLLRKLWMDWAEGLWKNRFTFVFLNVCENMGIA-ETSLLELGRDWPSEKIED 215
QY 255 LIRVPERLFIIDGPD-----ELKPSPHDPQGPWCLCWEERKRPTELLINSIRKLLPE 308
Db 216 IFSQPERILFIMDGPEQLKFNQLADLSD-----WRQOPMFIILSSLLQKKMLPE 268
QY 309 LSLLTTPALEKRLHLEHPRHVEILGFSEAEKKEYFYKYFHNAQAGQVNVVRONE 368
Db 269 SLLLTALGKLAQKHVFMRLRHPKLIKLLGFSESEKSYFSVFFGKSKALKVNFVRDNG 328
QY 369 PLFTMCFVPLVCVVCTCLOQLQEGGLLRQTSRTTAVYMLYLLSLMQPKGAPRLQPP 428
Db 329 PLFILCHNPFTCMLVCTCVKQRLERGEDLEINSQNTTYLYASFLITVF--KAGSQSFDPK 386
QY 429 PNOQ---RGLCSLAADGLANNQKILFEEQDLRKHGLDGEDVSAFLNMNIFOKDINCERYSP 485
Db 387 VNRARLKSALAAEGINTYTFVSHGDLRRNGLSESEGVMMVGMWELLORGDG---EAF 443
QY 486 IHLSEFQEPFAAMYIILDEGEAGDPD-QDVTRLLTYAFSERSFALATSRFLPGLNEET 544
Db 444 MHLCIOEFCARMFYLLKRPKDDPNPAIGSITQIVRASVVPQPTLLTQVGIFMGISTEEI 503
QY 545 RSHLEKSLCWKVSPIKMDLLQWIKSQASDGLTQOGLSEFSCLYEIQEEFIQOALS 604
Db 504 VSMLETSFGFPLSKDLKQBITQCLESLSQCEADREAIQFELFGLFETQKEFVTVKVMN 563
QY 605 HFQ--VIVVSNIAASKMEHWV--SSFCLKRCRQAQVHLHYGATYSADGEDRARCAGAHLL 661
Db 564 FFEVFIYIGNI-----EHLVIAFCLKHCHLTLRMCVENIFPDD---SGCISDYNEKL 616
QY 662 VOLRPERTVLLDAYSEHAAALCTPNLIELSLYRNALSGRVKLLCQGLRHPNCKLQNL 721
Db 617 V-----YWRLECSMFITKNFQILDMENTSLDDPSLAILCALQAQPVCKLRKL 664
QY 722 RLKRCRISSACEDLSAALIANKNLTFMDLSNGVGFPGMMLLCEGLRHPQCLQMLQLR 781
Db 665 IFTSVYFGHDS--ELFKAVLNPHLKLKLSLYGTSLSQSDIRHLICETLHPCKIEELILG 722
QY 782 KCOLESGACOCMAVSLGTNPHLVLDLTGNALDELGLRLLCQGLRHPVCRRLTLWLKICR 841
Db 723 KCDISSEVCEDIASVLACNSKLKHLISVENPLRDEGNTLLCEALKHSCALERLMLMVCC 782
QY 842 LTAACADELASTLSVNSQSLRELDLSNELDGLVILLCEGLRHPTCKLQTLRLGICLGS 901
Db 783 LTSVSCDSISEVLLCKSLSLDLSNALEDNGVASLCAALKHPGCSIRELWLMGCFLTS 842
QY 902 AACEGLSVVLQANHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLLDSCGLTAKAC 961
Db 843 DSCKDIAAVLTCNGKLTCLKLGHNEIGDGTGVRQLCAALQHPHCKLECLGLQTCFTRACC 902

```

QY 962 ENLYFTLGINOTLDTLVTNNALGDTGVRLCKRLSHPGCKLRVL 1006
Db 903 DDIAAALIACTLASLNDWLTALDADAVVLCALSHLDFDCAQL 947

RESULT 4
Q86W87 PRELIMINARY; PRT; 994 AA.
AC Q86W87;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL: BC050326; AAH50326.1;
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN_1.
DR SMART; SM00368; LRR_RI; 8.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
KW Hypothetical protein.
SQ SEQUENCE 994 AA; 113434 MW; 70B76PF836AC1E68 CRC64;

Query Match 26.5%; Score 1450.5; DB 4; Length 994;
Best Local Similarity 33.8%; Pred. No. 7.7e-117;
Matches 352; Conservative 197; Mismatches 403; Indels 89; Gaps 19;

QY 13 LSTLYLEBAVELKFKVLGVTAT-ELGEGKIPWGSMEKAGPLEMAQLLTHPGPREAR 71
Db 11 LAMYLEELKEEFKEKHEKHLKQNTLQELKQIPWTEVKASREELANLLIKHYEQQAN 70

QY 72 LALSTPERINRDLWREGQEDLVDPQETRYDYVRKPF-RLMEDRNARLGEVNLSHRY 130
Db 71 ITRIFQKMDRDLCKVMRER--TGTYTKYQAHAKQKTSRLWSSKSV----- 116

QY 131 TRLLLVKHSNPMVOQQLDTCRGHARTVGHQASPIKTIETLPEPDEPPEPRVTVMQG 190
Db 117 TEIHLVFEF-----EVKQEECD-----HLDRLFAP-KEAGKQPRVTIIQG 155

QY 191 AAGIGKSMIAHKVMDLWADGKLFQGRFDYLYFINCKRMQSATECSMDLIIFSCWPEPS 250
Db 156 PQGIGKTLMLKMAWSNKFIRDFRFLTYFFCCHELRB-LPPTSIALDIENEDPAA 214

QY 251 PLOELIRVPERLLFIIDGDELKPSHPDQGGPWCWCEKRPTELLNSLIRKLLPELS 310
Db 215 PITEIVSQPERLLFVIDFEELQGGNPEPDSLCGDLMEKRPVQVLLSLRLKMKLPEAS 274

QY 311 LLITRPTALEKHLLEHPRHVEILGFSEAEKVEFYKFNHAEAGQVFNVDNEPL 370
Db 275 LLIAIKPVCPELRDQVTEISYIQRGFNESDLVYFCFFKDPKRAEAFNLVRESEQL 334

QY 371 FMCVFLVVCVCTCLOQLEGGLRQTSRTTAVVYMLYLSLMQPKPGA--PRLQPP 428
Db 335 FSIQIPLLCWLCTLSLQEMQKGDALTCQSTTSVYSSFFVFNFTP-DGAEGTQQTQ 393

QY 429 PNQGLCSLADGLNQKLFEEQDLKRGHLDGEDVSAPLANWNIPOKINCERYTSFIHL 488
Db 394 HOLKALCSLAAGGMDTDFECEDDLRRNGVDADIPALLGTKILLKYGERESSYVFLHV 453

QY 489 SFQEFFAAMYI-----LDEGEGAGPDQDQVTRLLTEYAFSERSFLALTSRFLGLNNEET 544
Db 489 SFQEFFAAMYI-----LDEGEGAGPDQDQVTRLLTEYAFSERSFLALTSRFLGLNNEET 544

```

SQ SEQUENCE 1043 AA; 118910 MW; C321PBB65206DFF CRC64;
 Query Match 25.6%; Score 1401; DB 4; Length 1043;
 Best Local Similarity 32.8%; Pred. No. 1.7e-112;
 Matches 357; Conservative 171; Mismatches 405; Indels 154; Gaps 22;
 13 LSTYLELEAVELEKFKLYL-----GTATLGEKIPWGSMEKAGPLEMAQLLTHF 64
 17 LLPYLMALDQQLBEFKLEFPQQLMDFWSPAPQGHFRIPWANRAADPLNLSFLDEHF 76
 65 GPEAWRLALSTPERINKOLWERO-----REDL----- 94
 77 PKQAWKVLGIFQTMNLTSCVKRAEMKENVTOELQDPTQDELEMLEAAGNMQTQ 136
 95 VRDQETVYRDV-----RRKFLMEDRNARLGEV-VN 125
 137 CQDNQELDELETGNVQAGQDQNPQEBPEMLEEADHRRKYR--ENMKAELLETDWN 194
 126 LS-----HRYTRLLLVKHSNPMQVQQLDGTGRGHARTVGHQASPIKTIETLPEPDERPE 181
 195 ISWPKDHVYIRNTSKDEH-----EELQRLDPNRTAQA----- 228
 182 PPRIVMGAGAGIKSMIAHKVMDWADGKLFQGRFDYLVYINCREMNSQATSCSMODLI 241
 229 --QTVLVGRAGVGKTTLAMRAMLHWANGVLFQQRFSYVFLSCHKI-RYMKETTFABLI 285
 242 FSCWPEPSAPLOQLRVPERLFIIDGFDEL-----KPSFHDPOGPWCLCWEKRPTEL 295
 286 SLDWDFDAPTEEFMSQEKLLFIIDGFEIISRSRSESLDD--GSPCTDWMYQELPVTK 343
 296 LINSIRKKLPDELIIITRTALEKHLRLEHPRHVEILGPSAEKYEYKFFHNAE 355
 344 ILHSLKELVELPLATLITIKTFWVRDLKASLVNFCVQITGFTDGLRVYFMHFDSS 403
 356 QAGQVFNVRDNEPLTFWCFVPLVCMVCTCLOQLEGGLRQRTTAVYMLYLSL 415
 404 EYEKILQLRKNETLPHSCSAPMVCWVCSLQPKRVYDIQSTQITTSIYAFPSNL 463
 416 MQPK-----PGAPRLQPPNPGRLCSLAADGLNOKILFEBOJDLRKHGLDGEDVSA 466
 464 FSTAEDVLADDSWFG-----QWRALCSLAIEGLMSNFTFNKEDTEIEGLEVPFIDS 515
 467 FLNNMIFOKINCERYYSIHLSQFPPFAAMYILDEGE---GGAPDQDVTRLLTEYAF 523
 516 LYEFNLIQINDCGGCTTTHLSFQFFFAAMSFVLEPREFPFPHSTPKQEMKMLQHVLL 575
 524 SERSPLATSRFLGLNRETRSHLEKSLCWKVSPIHKMDLLQWIOSKQSGSTTLQOQS 583
 576 DKEAYWTPVLPVFFGLLNKNIARELEDTLHCKISPRYMEELLKWBELCKAESAIOFHI 635
 584 LBFFSCLYEIOBEFTQALSH-POVIVVSNIAASKMEHVSFCLKRCRQAQVHLHYGAT 642
 636 LRLFHCHLESQEDFTTKMLGRIFEVDL--NILEDEELQASSFCLKHKCKRL----- 684
 643 YSADGEDRARCAGATLVLQRPETVLLDAYS---EHLAALCTNPNIELSLYRNAL 699
 685 -----NKRLSVSSHILRELDLEILETSKFDPSRMHAWNSICSTLVNENHELDLSNKL 738
 700 GSRGVKLLCOGLRHPNCKLQNLKRCRISSACEDLSAALANKLTRMDLSGNGVGF 759
 739 HASSVKGCLCALNPKRCKVQKLTCKSV-TEPWVLDLIIALQNSKLTFLNFSNKLGMT 797
 760 GMLLCEGLRHPQCRLOMIQLRKQLESAGCQEMASVLTGNPHLVELDLTGNALEDGLR 819
 798 -VPLILKALRHSACNLKYLCKEKNLSAASQDLALFLTSIQHVTRICLGFNRLQDDGIK 856
 820 LLCQGLRHPVCRRLTLMKICRLTAACDELASTLSVNQSLRELDLSLNEGLDGLVLLC 879
 857 LLCAALTHPKCALERLELWFCQLAAPACKHLSDALQLQNRSLTHLNSKSLRDEGVKFLC 916
 880 EGRHPCTKLTQRLGICRLGSAACEGLSVVVLQANHLRELDLSFNDLGDWGLMLLAEG 939
 917 EALGRPDGNLQSLNLSGCSFTREGCGELANALSHNNVKILDLGENDLQDDGVKLLCEAL 976

QY 940 QHPACRLQKLDSCGLTAKACENLYFTLGINOTFLDLYLTNNALGDTGVRLLCRLSHP 999
 DB 977 K-PHRLHTLGLAKCNLTATACCQHLFVSLSSKSLVNLNLLGNELDTDGVKMLCKALKKS 1035
 QY 1000 GCKLRVL 1006
 DB 1036 TCRLQKL 1042
 RESULT 6
 Q86UB5 PRELIMINARY; PRT; 1375 AA.
 AC Q86UB5
 DT 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Similar to death effector filament-forming Ced-4-like apoptosis protein.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Strausberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051787; AAH51787.1; -;
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR GO; GO:0006952; P:defense response; IEA.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 DR SMART; SM00368; LRR_R1; 5.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR PROSITE 1375 AA; 154969 MW; 7C834D47BBD490FE CRC64;
 SQ SEQUENCE 1375 AA; 154969 MW; 7C834D47BBD490FE CRC64;
 Query Match 25.2%; Score 1377; DB 4; Length 1375;
 Best Local Similarity 33.9%; Pred. No. 3.1e-110;
 Matches 366; Conservative 136; Mismatches 357; Indels 218; Gaps 23;

QY 12 RLSTYLEAVELEKFKFLYLTA--TELGEKIPWMSMEKAGPLEMAQLLTHFGPEEA 69
Db 8 RLACYLEFLKEELKEFOLLANKAHSSSGETP-AQPEKTSMEVASYLVAQYGEQRA 66
QY 70 WRALSTFERINRDLWEROQR-----EDLVDPQBYTYDYV----- 106
Db 67 WDLALHTWEQWGLRSLCAQAEAGAGHSPSPYSPSEPHLGSPSPTSTAVLMPWIHELPA 126
QY 107 -----RRKFLMEDNARLGEVNLSHRYTRLLLVKHSNPMOVQ----- 146
Db 127 GCTQGSRRVRLQPLDTSGRWRREISASLLYQALPSSPDHESQSPNAPTSTAVILGSW 186
QY 147 ----- 146
Db 187 GSPFPQSLAPREQEAPCTQPLDTSGLIYTEIREREREKSEKGRPPMAAVGTPPOAHT 246
QY 147 -----QQLDTGRGHARTV----- 160
Db 247 SLOPHHHWPSPVRESLCSTWPKNEDFNQKFTQLLLQSRSHPRSQDPLVKRSWPDYVEE 306
QY 161 --GHOASPIKIELEFDEERPPRTVVMQGAIGKSMIAHKVWLDWADGKLFGQRFDP 218
Db 307 NRGLH-----IEIRDLFGELQDQ--PRVILQGAIGKSTLAKQVKEAWGRGOLYGRFQ 362
QY 219 YLFYINCREMNSQATECSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGDELKPSFHD 278
Db 363 HVFFVSCRELAQSKV-VSLAELIKDGTATPAPIRQILSRPERLLFIIDGDEPGWLOE 421
QY 279 PQGFWCLWEEKPTTELLNSLRKLLPELSLLITRPTALEKHLRLLHPRHVEILGF 338
Db 422 PSELCLHWSQOPADALLGSLGKTLPEASFLITARTTALQNLIPSLSEQARWVEVLGF 481
QY 339 SEAEKRYFYKYFHNAEAGOVNRYVNDNEPLFCVPLVWCVCTCLOQLEGGLLR 398
Db 482 SESSREKYFYRYTDERQAIRAFLVKSNEKELWALCLVWPVWSLACTCLMQQWKRREKUT 541
QY 399 QTSRTTAVMYLLSILMQPKAPRLQPPNQRGLCSLAADGLWQKILFEEDQLRKHG 458
Db 542 LTSKTTTTLCHVLAQALQAPLGPOL-----RDLCSLAEEGIWQKTLFSPDDLKRGH 595
QY 459 LGEDVSAFLNMWIFOKDINCERYYSFIHLSFOEFAAMYIILDEGGAGPQ-----DV 514
Db 596 LDGAILSTFLKMGIOEH-PIPLUSYFIHLCPFOEFAAMSYVL-EDEKGRGKHSNCIIDL 653
QY 515 TRLLTEYAFSERSFALTSRFLPGLNEETRSLSKSLCWKVSPIHKMDLLOWIQSKAQS 574
Db 654 EKTLEYAGI-HGLFGASTTFLGLLSDGEREMENIFHCRLSQ--GRNLMQVPSLQ-- 708
QY 575 DGTLOQGSLEPSCLYIEOEBEFTIOQALSHFOVIVVSNIAKMEHMSVSCFKRCRSAQ 634
Db 709 --LLLQPHSLCHLCLYETRNKTLQVMAHPEMGMC-VETDMELLVCTFCIKFSRHVK 765
QY 635 VLHLYGATYSADGEDRACSAHAHTLLVQLRPERTVLLDAYSHLAAALCTNPNLELSL 694
Db 766 KLOLI-----EGROHSTWSPVTVLFR--WVPVTDAYWOLFVSVLKVTRNLKELDL 815
QY 695 YRNALSGRVKLLCOGLRHFNCKQLNLRKRRISSACEDISAALIANKNLTRMDLSGN 754
Db 816 SCNSLSHSAVSKLTLRPRCLLETFLRAGCLTAEDCKDLAFLGRANQTLTDLSPN 875
QY 755 GVGFGFMMLLCEGLRHPCQLQMIQLRKQLESGACQEMASVLGTPNPHLVELDLTGNAL 814
Db 876 VLTDRAGHLCQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQONNLD 935
QY 815 DLGLRLCOGLRHVPCRLTLMKICRLTAACADELSTLSVNSQSLRELDLSNEL---- 870
Db 936 DVGVRLLCEGLRHPCAKLIRLGI-----DQTLSDMKRQELRALEQEKPOLIFS 985
QY 871 -GDGLVLLCEGL-----RHPTCKLQTLRLGICRLGSAACEGLSVVLOAHNLRELDLS 923
Db 986 RRPKPSVMTTEGLDTGEMSNSTSSLRQKRLGSEARAASHV-----AQAANKLLDVS 1035

RESULT 7
Q86W29 PRELIMINARY; PRT; 1200 AA.
ID Q86W29 AC Q86W29
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE NALP5.
GN NALP5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Techopp J., Martinon F., Burns K.;
RT "NALP5: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
DR EMBL; AY154460; AA018156.1; --
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR_2.
DR SMART; SM00368; LRR_R1; 11.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
SQ SEQUENCE 1200 AA; 134371 MW; 40C6DB1003B3B0A1 CRC64;

Query Match 24.2%; Score 1323; DB 4; Length 1200;
Best Local Similarity 31.2%; Pred. No. 1.3e-105;
Matches 354; Conservative 176; Mismatches 460; Indels 144; Gaps 21;

QY 17 LEELEAVELEKFKF-LYLGTATELGEGKIPWMSMEKAGPLEMAQLLTHFGPEAWRLALS 75
Db 69 LYELDEKEEFTFKELLKKSSSESTTCSIPOFEIENANVECLALLLHYYGASLAWATIS 128
QY 76 TFERINRDLWEEKQREDLV-----DPQTYRDRYVRKRFLL----- 112
Db 129 IFENMNLRTISEKA-RDDMKRHSPEDEATMTDQGPSKEKVPGISQAVQDSATAAETKE 187
QY 113 -----MEDNRALRGCVNLSHRYTRLLLVKHSNPMQVQOQLDTRGHARTVGHQASP 166
Db 188 QEISQAMBEQATAAB-----TEOEISQAMEQEGATAAETEEQ-----GGGDTWDYKSH 238
QY 167 IKIETLFEPEDEER-----PE-----PPRTVVMQGAAGIKSMIAHK 202
Db 239 VMTKPAEEDVRESFENTAADWPEMQTLAGAPDSRWGFRPRTVVLHGKSGIKSALARR 298
QY 203 VMLDWADGKLFGQRFYLYINCREMNSQATECSMODLIFSCWPEPSAPLOELIRVPERL 262
Db 299 IVLCAQGLYQGMFSYVFLPVREM-ORRKESSVTEFISREWPDSQAPVTEIMSRPERL 357
QY 263 LFTIDGDELKPSFHDQPGFWCLWEEKPTTELLNSLRKLLPELSLLITRPTALEK 322
Db 358 LFTIDGFDL-GLVNLNDTKLCKWAEKQPPFTLRSRLKLLPESFLIVTVRVDVGTET 416
QY 323 LHRLLHPRHVEILGFSEAEKRYFYKYFHNAEAGOVNRYVNDNEPLFCVPLVWCV 382
Db 417 LKSEVVSRYLLVRGISGEGORLHLLERGIHGEHQKTLQGLRAIMNNEELDLQCVPAVGS 476
QY 383 VCTCLOQLEGGLLRQTSRTTAVMYLLSILMQPKAPR---LQPPNQRGLCSLAA 439
Db 477 ICVALQQLDVGVESVAPFNQTLTGHAFAVPHQLTPRGVVRCLNLEERWILKRCFMAV 536
QY 440 DGLMNQKILFEEDQLRKHGLDGEDVSAFLNMWIFOKDINCERYYSFIHLSFOEFAAMY 499
Db 537 EGVWNRKSVFDDGLMVQGLGSESELRAFLHNMILLPDSHCEYYTFFHLSLODFCAALYY 596
QY 500 ILDEBEGGAGPQDVTLLTTEYAFSERSF-----LAUTSRPLFGLLNEETSHLE 549
Db 500 ILDEBEGGAGPQDVTLLTTEYAFSERSF-----LAUTSRPLFGLLNEETSHLE 549

```

597 VLE---GLEIPALCPPLYVEK--TKSMELKQAGFIHISLMWKRFLGLVSEDVRRPLE 650
550 KSLCWKVSPIKMDLLOWISKQSDGSLTQOGSLEFFSCLEYIETQEEFIIQALSHFQVI 609
651 VLLGCPVPLGVOKKLLHWISLLGQOPNATTPGDTLDAFHCLFQDKEFVRLALNSFQEV 710
610 VVSNIAKMBHMYSSFCCLKRC-----RSAQVL----- 636
711 WLP-INQNLDIASSFCLOHCPYLKIRVDVKGIFPRDESAEACPVVPLWMDKTLIEEQ 769
637 -----HL-----YGATYSAGEDRARCAGAH-TLLVQLRPERITVLDAYSEHL 679
770 WEDPCSMGLTHPHLRQLDGLSSILTERAMKTLCAKIRHPTCKIQTLMFRNAQITPGVQHL 829
680 AAALCTPNLELSIXRNALSGRGVLLCCGLRHPNCKLQNLRLKRCRISSSACEDLSAA 739
830 WRIVANRNRLSLNGGTHLKEEDVRVACEALKHPKCLLESRLDCCGLTHACYLKISQI 889
740 LIANKNLTRMDLSNGVGFPGWMLLCEGLRHPQCRLOLMQIRKCOLSGACQEMASVLTG 799
890 LTTSPSLKSLSLAGNKVTDQGVWPLSDALRVSCALQKLILEDGGITATGCGSLASALVS 949
800 NPHVELDLTGNALDGLRLCCGLRHPVCELRITLWLKICELTAACDELASTLSVNOS 859
950 NRSLSLHCLSNLSNGEVNLLCRSMRPHCSLQRLMLNQCHLDTAGCGFLALALMNSW 1009
860 LRELDLSLNEIGDLGVLLCEGLRHPCTCKIOTLRIGICRLGSAACGSLSVVQLQANHNL 919
1010 LTHLSLNMNPVEDNGVKLLCEVWRPESCHLODELVKLCHTAACCESLSVCISRSRHLKS 1069
920 LDLSFNDLGDWGLWLLAELGHPACRLQKMLDSCGLTAKACENIYFTLGINOTLTDLYL 979
1070 LDLTNLDGCGVAALCEGLKQKNSVLTGLKACGLTSDCCEALSLALSCNRHLTSLNL 1129
980 TNNALGDTGVRLLCKRLSHPGCKRLVL-----WFGMDLNKMTHTSLAALRVTKP 1029
1130 VONNFPKGMKLCFAFACPTSLQIIGLWKQVYPVQIRKL-----LEEVLQKLP 1179

RESULT 8
Q9BY26 PRELIMINARY; PRT; 287 AA.
AC Q9BY26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leucine-rich-repeat protein RNO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shami P.J., Kanai N., Wang L.Y., Vreeke T.M., Parker C.J.;
RT "Identification and characterization of a novel gene that is
RT upregulated in leukemia cells by nitric oxide.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231021; AAK14942.1; -
DR HSP; F13489; I44Y.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 1.
SQ SEQUENCE 287 AA; 31765 MW; BD3816C3255B2F9E CRC64;

Query Match 24.0%; Score 1314.5; DB 4; Length 287;
Best Local Similarity 80.1%; Pred. No. 8.4e-106;
Matches 261; Conservative 1; Mismatches 59; Gaps 2;

QY 710 GLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSNGVGFPGWMLLCEGLR 769
Db 21 GLLPQ--RRQLWLKRCRISSSACEDLSAALIANKNLTRMDLSNGVGFPGWMLLCEGLR 78
QY 770 HPQCRLOMQIRKCOLSGACQEMASVLTGNPHLVLDLTGNALDGLRLCCGLRHPV 829

597 VLE---GLEIPALCPPLYVEK--TKSMELKQAGFIHISLMWKRFLGLVSEDVRRPLE 138
550 KSLCWKVSPIKMDLLOWISKQSDGSLTQOGSLEFFSCLEYIETQEEFIIQALSHFQVI 889
651 VLLGCPVPLGVOKKLLHWISLLGQOPNATTPGDTLDAFHCLFQDKEFVRLALNSFQEV 198
610 VVSNIAKMBHMYSSFCCLKRC-----RSAQVL----- 202
711 WLP-INQNLDIASSFCLOHCPYLKIRVDVKGIFPRDESAEACPVVPLWMDKTLIEEQ 202
637 -----HL-----YGATYSAGEDRARCAGAH-TLLVQLRPERITVLDAYSEHL 1009
770 WEDPCSMGLTHPHLRQLDGLSSILTERAMKTLCAKIRHPTCKIQTLMFRNAQITPGVQHL 261
680 AAALCTPNLELSIXRNALSGRGVLLCCGLRHPNCKLQNLRLKRCRISSSACEDLSAA 1035
830 WRIVANRNRLSLNGGTHLKEEDVRVACEALKHPKCLLESRLDCCGLTHACYLKISQI 287
740 LIANKNLTRMDLSNGVGFPGWMLLCEGLRHPQCRLOLMQIRKCOLSGACQEMASVLTG 1035
890 LTTSPSLKSLSLAGNKVTDQGVWPLSDALRVSCALQKLILEDGGITATGCGSLASALVS 287
800 NPHVELDLTGNALDGLRLCCGLRHPVCELRITLWLKICELTAACDELASTLSVNOS 1035
950 NRSLSLHCLSNLSNGEVNLLCRSMRPHCSLQRLMLNQCHLDTAGCGFLALALMNSW 287
860 LRELDLSLNEIGDLGVLLCEGLRHPCTCKIOTLRIGICRLGSAACGSLSVVQLQANHNL 1035
1010 LTHLSLNMNPVEDNGVKLLCEVWRPESCHLODELVKLCHTAACCESLSVCISRSRHLKS 287
920 LDLSFNDLGDWGLWLLAELGHPACRLQKMLDSCGLTAKACENIYFTLGINOTLTDLYL 1035
1070 LDLTNLDGCGVAALCEGLKQKNSVLTGLKACGLTSDCCEALSLALSCNRHLTSLNL 287
980 TNNALGDTGVRLLCKRLSHPGCKRLVL-----WFGMDLNKMTHTSLAALRVTKP 1035
1130 VONNFPKGMKLCFAFACPTSLQIIGLWKQVYPVQIRKL-----LEEVLQKLP 287

RESULT 9
Q8BU40 PRELIMINARY; PRT; 982 AA.
AC Q8BU40;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to PAN2 protein.
GN E330028A19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK087843; BAC40024.1; -
DR MGI; MGI:2443697; E330028A19RIK.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFA2053585 CRC64;

Query Match 23.0%; Score 1259; DB 11; Length 982;
Best Local Similarity 31.0%; Pred. No. 3.6e-100;
Matches 325; Conservative 202; Mismatches 415; Indels 106; Gaps 24;

QY 13 LSTYLEEAEVLEKFKKLYL-GTATBELGEGKITPWGMEKAGPLEMAQLLITHEGPEAWR 71
Db 10 LMWYELNKKFVKFKFLKQVQLGLKQVSWTEVKKASQDQLASLLDKHYEEKPAWD 69
QY 72 LALSTERINRKDLWEGOREDLVRDPQETRYDYVRKFKRLMEDRNARLCECVNLSHRYT 131
Db 70 MTRFFQKINRKLIERAKKE--IDGCPKLYRAHMTK-----MTDSS 111
QY 132 RLLLVKSHSNPMVQVQQLDTRGHARTVGHQASPIK-----KFTDDYDCFNLFQ-SKGTESKP 183
Db 112 RAFTI-----SIQNFKE-----KFTDDYDCFNLFQ-SKGTESKP 147
QY 184 RTVMQGAAGIGKSMIAHKNLMDWADOKLFGQRPDYLYFYINCREMNSQATECSMDLIFS 243

```

Db 148 QVFLSGAGVGKTLMLKRLMLAWIESPVFLHFKFSYIFFCREVBQKLT-ASLAELISR 206
 Qy 244 CWPEPSAPLOELIRVPERLFIIDGFDLKPSPHDQGPWCLCWEEKRPTLLNSLRK 303
 Db 207 EWGSPAPIEILSKPKLLFIIDSLGMECDLFWKSELCDNTEKQPVNVLSSLLR 266
 Qy 304 KLPELSLLITRPTALEKHLRLEHPRHVEIL-GFSEAEKKEYFYKYFHNAEQGVFN 362
 Db 267 KMLPESSLLISATPESPEKMEINREY-THVKIIGLKAKERNIKMSFHLRFDQNEAFAFS 325
 Qy 363 YVRDNEPLFMVCFVPLVWVCTCLOOQEGGLLROTSRTTAVMYLLSLMQPFG- 421
 Db 326 LVRENEQLFVFCVPLVWVATCBEIEKGRDPVSCICRTTSLYTHIFNFIQVAH 385
 Qy 422 APRLQPPNPORGCLSLAAGLWNGKILFEEDQDLKHLGDGEDVSAPLNWIFOKDINCR 481
 Db 386 SPKSKSDQQLGLCSLAEGWNTDTFVFGELARRNGIMSDPILLDGLMINIRESEK 445
 Qy 482 YGFIHLSFORFFAAMYIILDEGGGAGPDQDVTLLTEYAFSERSFLALTSR----- 534
 Db 446 SYIFLHPSVQEVCAAFYLLKSHVD--HPSQEVKSI-----EKLMAFLKVKVQWIF 496
 Qy 535 ---FLGLNEETSHLEKSLWKVSPHINKDILLOWIQSKAOSDGTLOQGSLEFFSCLY 591
 Db 497 FGSFIFGLHESQKLEAFHGHLSQEIIRQLYQCLLETISGNEELQEQIDGMKLFYCLF 556
 Qy 592 EIQEEFFIQALSHFOVIVVSNLASK--MEHMYSSFCLEKRCRAOVHLHYGATYSADGD 649
 Db 557 EMDDDTFLVEAMCMQI---NFVAKDYSVIVVAHCLKHCFTLKLKF-----STQ 605
 Qy 650 RARCSAGATLLVOLRPERVTLDDAYSEHIAAALCTNPNLIELSLYRNALSGRGVLLCQ 709
 Db 606 NVLSGAQESYM-----ER--LLTAWN-HICSVFIISKDQELRMKDTNLSGSAFSLYN 657
 Qy 710 GLRHPNCKLQNLKRCRISSS---ACED--LSAALTANKULTMDLSGNGVGPFGMMLLC 765
 Db 658 NLKHYNTLVN-----VANNVFFCEKYLFFELIQNCNLOHNLSTILSHSDVKLL 711
 Qy 766 EGLRHPQRLQMTQLRKCOLESACOMASVLGTNPHLVELDLTGNALEDGLRLLCGL 825
 Db 712 DVLQAECNEELVVAACISSDCKVFASVLISNKTILKHLNLASNTL-DKGIASLCKSL 770
 Qy 826 RHPVCLRLTWLKI CRLTAAACDELASTLSVNGSRELDLSNDELGDGLVLLCEGRHP 885
 Db 771 CHPDCTLEHLVLANCSINEKWDYLSVLRRNKTLSHLDISSLNDKDEGLKVLCRALTLP 830
 Qy 886 TKLOTLRLGICRLGSAACEGLSVLQANHLRELDSLNDLGDWGLMLLAEGLOHPACR 945
 Db 831 DSVLSLSLRHCLITISGQDLAEVLRRNNQNLVSLQVSNKKLEDTGVLLCDKAIRPNCH 890
 Qy 946 LOKLMLDSOGLTAKACENLYFTLGINQTLTDLTYLTNNALGDTGVRLCKRLSHPGCKLRV 1005
 Db 891 LEDLGLAECLATGACEDLASTFTQCKTLWANNLLKNALDYNGLVLCALQOICATVY 950
 Qy 1006 LMLFGMDLNKMTSHRLAALRVTKPYLDI 1033
 Db 951 IGLQITDFTDTQAFVAEQEKNPCLRI 978

RESULT 10

Q86W28
 AC Q86W28 PRELIMINARY; PRT: 1029 AA.
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE NALP8.
 GN NALP8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22451042; PubMed=12563287;
 RA "Schopp J., Martinon F., Burns K.;
 RT "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
 DR EMBL; AY154463; AA018159.1; --
 DR InterPro; IPR001611; LRR
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_RI; 7.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR PROSITE; PS00837; NACHT; 1.
 SQ SEQUENCE 1029 AA; 117335 MW; 7F8CEC14577303BA CRC64;

Query Match 22.2%; Score 1216.5; DB 4; Length 1029;
 Best Local Similarity 33.1%; Pred. No. 1.9e-96;
 Matches 325; Conservative 165; Mismatches 391; Indels 101; Gaps 20;

Qy 16 YLEBEAVALKPKKYLGTATLGEKIP--WGSMEKAGPLEMAQLLITHFGEAAWRLA 73
 Db 44 YMRNVSHLEELQRFKQLL--LTSLSTGMPITWQVETASWAEVVLHLLIERPGRADVT 101
 Qy 74 LSTFERINKDLNERQEDLVDRDPOETVYRKYRFRMLMEDRNALGEC-VNLS----- 127
 Db 102 SNIFAIMNCDKMCVVVRE-----INAILPTLEPEDLVNGETQVNLGEGSG 148
 Qy 128 --HRYTRLLLVKHSNPMQVQQQLLD--TCRHARTVGHQA---SPIKITLPEPEERP 180
 Db 149 KIRY-----KSNVNEKFFPDIITWPGNQRFYQGVHRHEEYFCLLLKRPQG 200
 Qy 181 EPRTYVMQGAAGIGKSLAHKVMWDADGKLFQGRFDYLFYINCRMMQSAATECSMODL 240
 Db 201 RQPKTVAIGAGIGKTLAKKVMFEARNKFVAKWCAFYHCEVQNT-TDQSFSEL 259
 Qy 241 IFSCWEPSPAPIQLIRVPERLFIIDGDELKPSFHDQGPWCLCWEEKRPTLLNSL 300
 Db 260 IEQWPGSQDLVSKMSKPDQLLLDGFELTSTLIDRLDLSEDRWKQLPGSVLLSSL 319
 Qy 301 IRKKLIPELSLLITRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKYFHNAEQGV 360
 Db 320 LSKTMLPEATLLIMRFTSWQTCCKLLKCPSLVTLPGFNIMEKIKTFQMYFGHTEGQV 379
 Qy 361 FNYVRDNEPLFTMCPVPLVWVCTCLOOQEGGLLROTSRTTAVMYLLSLMQPXP 420
 Db 380 LSFAMENTILFMCVRPVVWVWVCSGLKQOMERGNLITQCPNATSVFVYISSLFPTR- 438
 Qy 421 GAPRLQPPNQ--RGLCSLAADGLWNGKILFEEDQDLKHLGDGEDVSAPLNWIFOKDI 477
 Db 439 -AENSFKHQALQELCHLAADSMWRKWLKEDLEAKLDQGTAVTAFGLMSILRITA 497
 Qy 478 NCERYYSFIHLSFQBFFAAMYIILDEGGGAGPDQ-----DVTRELLTEYAFSERS 527
 Db 498 GEEDHYVTLTVTFQBFFAALFYV-----CFQRLKNPHVLVSHVNIQLIASPRGS-KS 550
 Qy 528 FLALTSRFLGLNNEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQSDGTLOQGSLEFF 587
 Db 551 YLSHMGLEFLFGFLNEACASAVEQSPQCKVSFGNKRKLLKVIPLLHKCDPPSPSGVPLF 610
 Qy 588 SCLYEIOEEFFIQALSHFOVIVVSNLASKMEHMYSSFCLEKRCRAOVHLHYGATYSADG 647
 Db 611 YCLHIREAFVSQALNDYHKKVVL-RIGNNKEVQVSFCLKRC---QYLHEVELVTNLNF 666
 Qy 648 EDRARCSAGATLLVOLRPERVTLDDAYSEHIAAALCTNPNLIELSLYRNALSGRGVLL 707
 Db 667 MNVWKLSSSSHP--GSEAPESNG-LHRWQDLCSVFATNDKLVLTMTNSVLGPFPLKAL 723
 Qy 708 CQGLRHPNCKLQNLKRCRISSSACEDLSAALTANKULTMDLSGNGVGPFGMMLLCBG 767
 Db 724 AALRHPQCKLQKLLLRVN-STMLNQDLIGVLTGNQHLRYLEIQHVEVESKAVKLCRV 782


```

QY 768 LRHPQCRIQMIQRLKQESGACQEMASVLTGNPH----- 802
Db 783 LRSFRCRLQCLREDCIATPRTWIDGNNLQNGHKLTLILRKNLENGAYVLSVAQLE 842
QY 803 -----LVELDITGNALEDGLRLLCGLRHPVCRRLTLMKICRTATAACDELASTLS 855
Db 843 RLSQSKMLTHLSABNALKDGAHWNALPHLRCPQLRLVLRKCDLTFNCCQDMISALC 902
QY 856 VNQSIRELDLSNELGDLVLLCEGLRHPCKQLQTLRLGICRLGSAACEGLSVVLQANH 915
Db 903 KNTKLSLDLSFNSKDGVIILCEALKNPCTLIQILENCLFTISICQAWASMLRKQ 962
QY 916 NLRELDLSFNDLGDWGLWLLAE 937
Db 963 HLRHLDLSKNAIGVYGILTCE 984

RESULT 11
ID Q7TPU9 PRELIMINARY; PRT; 1111 AA.
AC Q7TPU9;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Maternal effect gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053384; AAH53384.1; -
SQ SEQUENCE 1111 AA; 125429 MW; A1ACBE8E98A546F8 CRC64;

Query Match 21.4%; Score 1168.5; DB 11; Length 1111;
Best Local Similarity 31.5%; Pred. No. 3.3e-92;
Matches 294; Conservative 165; Mismatches 406; Indels 69; Gaps 13;

QY 162 HQASP-IK-ITLTFEPPERPEPRIVMCGAGICKSLAHKVMLDWADGKLFQGRFDY 219
Db 168 HYDSPENKLSDAFKP-YKQTFQPHITLHGRPGVGKSLARSIVLGNAGQKLFQ-KMSF 225
QY 220 LFVINCENMQSATECSMODLPSFCWPEAPQLIRVPERLLFIIDGFDLKPSPHDP 279
Db 226 VIFPSVREIKWT-EKSLAQIAKECPDSWDLTKMSQPERLLFVIDGLDDMSVLQHD 284

```

```

QY 280 QGPCLCWEKRPTELLNSLIRKKLPELSLITTRPTALEKLRLLEHPRHVEILGFS 339
Db 285 DMTLSRDWKDEQPTIYILMYSLLRKALLPOSFLITTRNTGLEKLSKMSVVSPLYILVEGLS 344
QY 340 EABEKEYFYKFNHAEQAGOVNVRDNEPELFTWCFVPLVCWVCTCIQQOQLEGGLLRQ 399
Db 345 ASRSQVLVLEINISBESBRIQVHSLIENHQLFDQCQSPVCSINCEALQOKLKGKCTL 404
QY 400 TSRTTAVY---MAYLLSLMQPKPGAPRLQPPPNORGLCSLAADGLNNQKILFEQDLRK 456
Db 405 PCQTLTGLYATLVFHLTKRPSQSALSOEQITLVGLCMAAEGVWTRSVFYVDDDLKN 464
QY 457 HGLDGEDVSAFLNMI-FKDINCERYYSFTHLSQEPFAAWYLLDSEGGAGDQ--- 512
Db 465 YSLKESEILALFHMNILLQVGHNSEQCQVFSHLSLQDFFAALYYVL---EGLEGWNQHC 521
QY 513 --DVTLLITVEAFSERSFLALTSRFLGLNEETRSHLEKSLCKWKSVPSPHIMDLILQWIS 570
Db 522 FIENQRSIMEVKRTDDTFLGLMKRFLGLMKNKDIILKTLVLFVFPVITVEQKLQHWVSL 581
QY 571 KAQSDGSTLQOQSLEFFSCLYIEOEERFIOALSHFQVIVVSNIASKMHVMSFCLKRC 630
Db 582 IAAQVNGTSPMDTLDAFYCLFESODEEFVGALKRFQEWLL-INQKMDLVKSSYCLKHC 640
QY 631 RSAQVHLH-----YCATYSADGE----- 648
Db 641 QNLKAIRVDIRDLSVDNTLCPVVTVOETQCKPLMEWNGFCSVLGSRLNKELDILG 700
QY 649 DRARCSAGAHLLVQLRPE-----RTVLLDAYSEHLAAALCTNPNIELSILYRNAL 699
Db 701 DSILSQRAMKILCLELRNQSCKIQKLTFSAEVVSGLKHLWKLLFSNQNLKYLNLGNTPM 760
QY 700 SGRGVKLLCQGLRHPNCKLQNLRLKRISSACEDLSAALIANKNLRTMDLSGNGVGP 759
Db 761 KDDMKLACEALKHPKCSVETLRDSCBLTIIGYEMISTLIISTTRKCLSLAKNRVGK 820
QY 760 GNMILCEGLRHPQCELOMIQRLKCOLESGACQEMASVLTGNPHLVELDITGNALEDGLR 819
Db 821 SMISLGNALSSMCLLQKLIIDNCGLTSPASCHLIVSALFSNQNLTHLCLSNNSLGTGVQ 880
QY 820 LILCQGLRHPVCRRLTLMKICRTATAACDELASTLSVNQSLRELDLSNELGDLVLLLC 879
Db 881 QLCQFLRNPICALQRLILNHNCNIVDDAYCFLAMRLANNTKLTHLSLTWNVPVGDGAMKLLC 940
QY 880 EGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQAHNHLRELDLSFNDLGDWGLWLLAEGL 939
Db 941 EALKEPTCYQLLELVDCQLTQNCCEDLACMTITTTKHLKSLDLGNALGDKGVITLCEGL 1000
QY 940 QHPACELQKWLDSCLGTAKACENLYFTLGINOTLTDLVLTNNALGDTGVRLICKRLSHP 999
Db 1001 KQSSSLRRKLGUGACKLITSNCCCEALSLAISCNPHLSNLNVKNDPSTSGMLKLSAQFQ 1060
QY 1000 GCKLRVLMFLGMDLNKMTSHSLAALRVTKPYLDI 1033
Db 1061 VSNLGIIGLWKQEYVARVERQLEEFVFKPHVI 1094

RESULT 12
ID Q95LZ7 PRELIMINARY; PRT; 713 AA.
AC Q95LZ7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:56:27 ; Search time 58 Seconds
(without alignments)
5042.012 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 1035
Sequence: 1 MLRTAGDGLCLSTYLEEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1035	100.0	1035	5 ADE36417	Human PAA
2	1019	98.5	1099	5 AAO17857	Pyrin dom
3	818	79.0	1061	5 AAO15590	Human PYR
4	661	63.9	865	7 ADC31287	Human nov
5	603	58.3	603	5 ADE36457	Human PAN
6	445	43.0	565	5 ABG97475	Human nuc
7	408	39.4	521	6 ABU99120	Novel hum
8	350	33.8	582	6 ABU99119	Novel hum
9	254	24.5	449	5 ABB77910	Amino aci
10	217	21.0	287	5 AAE21064	Novel hum
11	217	21.0	344	5 AAE21062	Novel hum
12	190	18.4	190	5 ADE36456	Human PAN
13	190	18.4	137	7 ADC32968	Human nov
14	172	16.6	287	5 AAE21063	Novel hum
15	126	12.2	168	4 ABG04570	Novel hum
16	94	9.1	94	4 AAM78275	Human bon
17	81	7.8	159	3 AAB42357	Human ORF
18	75	7.2	75	5 ADE36399	Human PAA
19	56	5.4	56	4 AAM78296	Human bon
20	49	4.7	65	5 AAO15589	Human PYR
21	44	4.3	89	4 AAM83229	Human imm
22	29	2.8	29	5 AAE21075	Novel hum
23	29	2.8	29	5 AAE21074	Novel hum
24	29	2.8	29	5 AAE21072	Novel hum
25	28	2.7	28	5 AAE21071	Novel hum

26	28	2.7	28	5 AAE21068	Novel hum
27	28	2.7	28	5 AAE21069	Novel hum
28	28	2.7	28	5 AAE21070	Novel hum
29	26	2.5	29	5 AAE21073	Novel hum
30	22	2.1	28	5 AAE21067	Novel hum
31	16	1.5	29	5 AAE21076	Novel hum
32	13	1.3	13	5 ABB77911	Protein k
33	13	1.3	13	5 ABB77913	Protein k
34	13	1.3	13	5 ABB77912	Protein k
35	13	1.3	75	5 AAO21931	PAAD doma
36	13	1.3	518	5 AAO17871	Novel hum
37	13	1.3	719	5 ADE36452	Human PAA
38	13	1.3	1034	4 AAE07514	Human PYR
39	13	1.3	1034	6 ABU08503	Human PYR
40	13	1.3	1034	6 ABU63315	Human PYR
41	12	1.2	240	4 ABG04098	Novel hum
42	12	1.2	385	5 AAU79523	Human MAT
43	12	1.2	385	6 AAE31746	Human MAT
44	12	1.2	1033	5 ABU65214	Human NOV
45	12	1.2	1143	6 ADA45221	Human MAT

ALIGNMENTS

RESULT 1

ADE36417
ID ADE36417 standard; protein; 1035 AA.

XX AC ADE36417;

XX DT 29-JAN-2004 (first entry)

XX DE Human PAAD and nucleotide binding protein PAN6.

XX KY cytosolic; immunosuppressive; vulnarary; antiinflammatory; vasotropic;
XX KY antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
XX KY antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
XX KY NFkappaB activation inhibitor; PAAD domain containing polypeptide;
XX KY PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
XX KY apoptosis-associated speck-like protein; caspase recruitment domain 2;
XX KY ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
XX KY apoptosis; NFkappaB induction; cytokine processing;
XX KY cytokine receptor signaling caspase-mediated proteolysis;
XX KY c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
XX KY inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
XX KY neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
XX KY fibrosis; smooth muscle cell proliferation; balloon angioplasty;
XX KY restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
XX KY arthritis; lupus; schrojen's syndrome; Crohn's disease;
XX KY ulcerative colitis; graft versus host disease; stroke; heart failure;
XX KY neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
XX KY cancer therapy; PAAD domain family; human; PAN6.

OS Homo sapiens.

XX US2003077699-A1.

XX 24-APR-2003.

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

XX 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (CHUZ/) CHU Z.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (ARIZ/) ARIZA M E.

XX (STEH/) STEHLIK C.

PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;
 DR WPI; 2002-471256/50.
 DR N-PSDB; ADE36416.
 XX
 XX Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 XX Claim 18; SEQ ID NO 24; 93pp; English.
 XX
 XX The invention describes an isolated PAAD domain containing polypeptide
 CC comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NF-kappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6.
 XX
 XX Sequence 1035 AA;

Query Match 100.0%; Score 1035; DB 5; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRTAGRDGICRLSTYLEELAEVLEKFKLYLGATLGECKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGRDGICRLSTYLEELAEVLEKFKLYLGATLGECKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEAWRLALSTFERINRKLWGRQREDLVDPQETRYDYVRRKFLMEDRNARL 120
 DB 61 ITHFGPEAWRLALSTFERINRKLWGRQREDLVDPQETRYDYVRRKFLMEDRNARL 120
 QY 121 GECVNLSHRYTRLLVKEHSPMVQVQQLDTGCHARTVGHQASPIKTIETLFPDDERP 180
 DB 121 GECVNLSHRYTRLLVKEHSPMVQVQQLDTGCHARTVGHQASPIKTIETLFPDDERP 180
 QY 181 EPPRTVMQGAAGIKGSLAHKVMWLDWADGKLFQGRFDYLYINCREMNQATSCSQDQL 240
 DB 181 EPPRTVMQGAAGIKGSLAHKVMWLDWADGKLFQGRFDYLYINCREMNQATSCSQDQL 240
 QY 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGWCLCWEEKRPTLLNSL 300
 DB 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGWCLCWEEKRPTLLNSL 300
 QY 301 IRKLLPELSLITTRTALEKHLRLEHPRHVEILGFSEAEKKEYFYKYPHNAEQAGV 360
 DB 301 IRKLLPELSLITTRTALEKHLRLEHPRHVEILGFSEAEKKEYFYKYPHNAEQAGV 360
 QY 361 FNYVRDNEPLFTMCFVPLVCVWVCTCQQQLGEGGLLRQTSRTTAVYMLYLLSLMOPKP 420
 DB 361 FNYVRDNEPLFTMCFVPLVCVWVCTCQQQLGEGGLLRQTSRTTAVYMLYLLSLMOPKP 420
 QY 421 GAPRLQPPPNORGLCSLAADGLMNQKILFEEDQLRKHGLDGEDVSAFLNMNIFOKDINCE 480

DB 421 GAPRLQPPPNORGLCSLAADGLMNQKILFEEDQLRKHGLDGEDVSAFLNMNIFOKDINCE 480
 QY 481 RYYSFHLSPQBFPAAMYIILDEGGAGAGPDQVTRLLITAYASERSFALTSRFLGILL 540
 DB 481 RYYSFHLSPQBFPAAMYIILDEGGAGAGPDQVTRLLITAYASERSFALTSRFLGILL 540
 QY 541 NEETRSHLEKSLCWKSPHIMKMDLLOWIQSKAOSDGTSTLQOQSLSEFFSYLQEEBFQ 600
 DB 541 NEETRSHLEKSLCWKSPHIMKMDLLOWIQSKAOSDGTSTLQOQSLSEFFSYLQEEBFQ 600
 QY 601 QALSHFOVIVVSNIAKMEHWSFCLKRCRQAQVLYHLYGATYSADGEDRARCAGAHTL 660
 DB 601 QALSHFOVIVVSNIAKMEHWSFCLKRCRQAQVLYHLYGATYSADGEDRARCAGAHTL 660
 QY 661 LVQLRPERTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRGVLLKCOGLRHPNCKLON 720
 DB 661 LVQLRPERTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRGVLLKCOGLRHPNCKLON 720
 QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSNGVGFPGWMLICEGLRHPQCRLOMIQL 780
 DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSNGVGFPGWMLICEGLRHPQCRLOMIQL 780
 QY 781 RKQLESAGCAQCEMASVLGTNPVHVELDTGNALDELGLRLCCOGLRHPVCLRTLWLKIC 840
 DB 781 RKQLESAGCAQCEMASVLGTNPVHVELDTGNALDELGLRLCCOGLRHPVCLRTLWLKIC 840
 QY 841 RLTAACDELASTLSTVNSQSLRELDLSNELGDLGVLLICEGLRHPCTCKLOTLRIGICRLG 900
 DB 841 RLTAACDELASTLSTVNSQSLRELDLSNELGDLGVLLICEGLRHPCTCKLOTLRIGICRLG 900
 QY 901 SAACEGLSVVLOAHNRLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKWLDSGLTAKA 960
 DB 901 SAACEGLSVVLOAHNRLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKWLDSGLTAKA 960
 QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCCKRLSHPGCKLRVLMFGMDLNKMTSHR 1020
 DB 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCCKRLSHPGCKLRVLMFGMDLNKMTSHR 1020
 QY 1021 LAALRVTKPYLDIGC 1035
 DB 1021 LAALRVTKPYLDIGC 1035

RESULT 2
 AAO17857
 ID AAO17857 standard; protein; 1099 AA.
 XX
 AC AAO17857;
 XX
 XX 20-AUG-2002 (first entry)
 XX
 DE Pyrin domain containing protein NALP3/PY5-hs.
 XX
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteoparitic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX
 OS Unidentified.
 XX
 XX W0200240668-A2.
 XX
 XX 23-MAY-2002.
 XX
 XX 30-OCT-2001; 2001WO-EP012545.
 XX
 XX 15-NOV-2000; 2000DE-01056687.
 XX
 XX 30-NOV-2000; 2000DE-01059595.
 XX

BA (APOT-) APOTECH RES & DEV LTD.
 XX PI Tschoop J, Martinon F;
 XX DR WPI; 2002-427093/45.
 XX DR N-PSDB; AAL47129.
 XX PT New DNA encoding protein with pyrin domain, useful for treating diseases
 PT involving impaired signal transduction, particularly inflammation, also
 PT proteins and antibodies.
 XX PS Claim 5; Fig 1; 116pp; German.
 XX CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention
 XX CC Sequence 1099 AA;
 SQ

Query Match 98.5%; Score 1019; DB 5; Length 1099;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRTAGDGLCRISTYLEEAEVLEKFKLYLTATLGEKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGDGLCRISTYLEEAEVLEKFKLYLTATLGEKIPWGSMEKAGPLEMAQLL 60

QY 61 ITHFGPEAWRLALSTFERINRDLWERGQEDLVDPQETRYDYVRKPRFIMEDRNARL 120
 DB 61 ITHFGPEAWRLALSTFERINRDLWERGQEDLVDPQETRYDYVRKPRFIMEDRNARL 120

QY 121 GECVNLSHRYTRILLVKEHNPQVQOQLLDTRGHARTVGHQASPIKTIETLPEDEEP 180
 DB 121 GECVNLSHRYTRILLVKEHNPQVQOQLLDTRGHARTVGHQASPIKTIETLPEDEEP 180

QY 181 EPRTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFDYLYINCREMNSQATSCSMDL 240
 DB 181 EPRTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFDYLYINCREMNSQATSCSMDL 240

QY 241 IFCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGPWCLCWEKRPTELLINSL 300
 DB 241 IFCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGPWCLCWEKRPTELLINSL 300

QY 301 IRKKLPELSLLITTRPTALEKHLRLLEHPRHVEILGFSAEKREYFYKVFHNAEQGV 360
 DB 301 IRKKLPELSLLITTRPTALEKHLRLLEHPRHVEILGFSAEKREYFYKVFHNAEQGV 360

QY 361 FNYVRDNEPLFTMCFVPLVCWVCTCLQQQLEGGLLRQTSRTTAVYMLYLLSLMQPKP 420
 DB 361 FNYVRDNEPLFTMCFVPLVCWVCTCLQQQLEGGLLRQTSRTTAVYMLYLLSLMQPKP 420

QY 421 GAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHGLDGEDVSAFLNNIFQKQINCE 480
 DB 421 GAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHGLDGEDVSAFLNNIFQKQINCE 480

QY 481 RYYSFTHLSFOERFAAMYVILDEGEGAGDPQDVTRLLTYAFSESRFLATSRFLGLL 540
 DB 481 RYYSFTHLSFOERFAAMYVILDEGEGAGDPQDVTRLLTYAFSESRFLATSRFLGLL 540

QY 541 NEETRSHLEKSLCWKVSPIHKMDLLOWIQSKASDGLTQQSGLEFFSCLYEIQEEFFIQ 600
 DB 541 NEETRSHLEKSLCWKVSPIHKMDLLOWIQSKASDGLTQQSGLEFFSCLYEIQEEFFIQ 600

QY 601 QALSHFOVIVVNSIASKMEHNVSSFCIKRCRQAQVHLHYGATYSADGEDRARCAGAHTL 660
 DB 601 QALSHFOVIVVNSIASKMEHNVSSFCIKRCRQAQVHLHYGATYSADGEDRARCAGAHTL 660

QY 661 LVQLRPERTVLLDAYSEHIAAALCTNPNIELSLYRNALSGRGVKLLCOGLRHPNCKLQ 720
 DB 661 LVQLRPERTVLLDAYSEHIAAALCTNPNIELSLYRNALSGRGVKLLCOGLRHPNCKLQ 720

QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQL 780
 DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQL 780

QY 781 RKQLESAGCOEWASVLTGNPHLVDELDTGNALEDLGLRLCOGLRHPVCRLETLWKIC 840
 DB 781 RKQLESAGCOEWASVLTGNPHLVDELDTGNALEDLGLRLCOGLRHPVCRLETLWKIC 840

QY 841 RLTAACDELASTLSVNSQISRLDELDSLNEGLDGLVLLCEGLRHPCKLQTLRLGICRLG 900
 DB 841 RLTAACDELASTLSVNSQISRLDELDSLNEGLDGLVLLCEGLRHPCKLQTLRLGICRLG 900

QY 901 SAACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKA 960
 DB 901 SAACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKA 960

QY 961 CENLYFTLGINQTLTDLTLTNALGDTGVRLCKRLSHPCCKLRVLWLFQMDLNKMTWS 1019
 DB 961 CENLYFTLGINQTLTDLTLTNALGDTGVRLCKRLSHPCCKLRVLWLFQMDLNKMTWS 1019

RESULT 3
 AAO15590
 ID AAO15590 standard; protein; 1061 AA.
 XX AC AAO15590;
 XX AC AC
 XX DT 31-OCT-2002 (first entry)
 XX DE Human PYRIN-8 protein #2.
 XX DE Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-8.
 XX OS Homo sapiens.
 XX PN WO200261049-A2.
 XX PD 08-AUG-2002.
 XX PF 31-JAN-2002; 2002WO-US002967.
 XX PR 31-JAN-2001; 2001US-0265231P.
 PR 10-SEP-2001; 2001US-0318645P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA (AMHP) WYETH.
 XX Bertin J, Wang W, Blatcher M;
 XX WPI; 2002-627477/67.
 DR N-PSDB; AAL44363.
 XX New PYRIN polypeptides and nucleic acids useful for modulating and
 PT diagnosing stress-related, apoptotic and inflammatory responses, or for
 PT treating inflammatory and immune system disorders, cancers, or
 XX neurological diseases.
 PS Claim 8; Fig 8; 167pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 CC useful for modulating and diagnosing stress-related, apoptotic and
 CC inflammatory responses. The PYRIN protein and DNA sequences are useful

CC for treating: inflammatory disorders and immune system disorders (e.g.
 CC Crohn's disease, reactive arthritis, multiple sclerosis, contact
 CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
 CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
 CC (e.g. systemic lupus erythematosus and arthritis); and neurological
 CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 CC protein and DNA sequences may also be used in screening assays, detection
 CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 CC predictive medicine (e.g. diagnostic assays, clinical trials and
 CC pharmacogenomics) and transcription profiling. The present amino acid
 CC sequence represents a human PYRIN-8 protein
 XX
 SQ Sequence 1061 AA;

Query Match 79.0%; Score 818; DB 5; Length 1061;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 938; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 97 DPQTYRDYVRKPELMEDRNARLGEVNLNHRVYTRLLLVKEHNSPMQVQQLDTRGH 156
 Db 124 DPQTYRDYVRKPELMEDRNARLGEVNLNHRVYTRLLLVKEHNSPMQVQQLDTRGH 183
 QY 157 ARTVGHQASPIKIEITLPEDEERPEPRVTVMQGAAGIGKSMIAHKVMDWADKLFQGR 216
 Db 184 ARTVGHQASPIKIEITLPEDEERPEPRVTVMQGAAGIGKSMIAHKVMDWADKLFQGR 243
 QY 217 FDYLFYINCRMNQATSCMODLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPFS 276
 Db 244 FDYLFYINCRMNQATSCMODLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPFS 303
 QY 277 HDQPGWCLCWEERKPELLNLSLRKLLPELSLITRPTALEKLRHLEHPRHVEIL 336
 Db 304 HDQPGWCLCWEERKPELLNLSLRKLLPELSLITRPTALEKLRHLEHPRHVEIL 363
 QY 337 GFSEARKEFYKYFHNAQAGQVFNVRDNEPLFTMCVFPLVCMVCTCLOQLLEGGL 396
 Db 364 GFSEARKEFYKYFHNAQAGQVFNVRDNEPLFTMCVFPLVCMVCTCLOQLLEGGL 423
 QY 397 LRQTSRTTAVMYLILSLMOPKPGAPLQPPNQRGLCSLAADGLWNOKILFEEQDLRK 456
 Db 424 LRQTSRTTAVMYLILSLMOPKPGAPLQPPNQRGLCSLAADGLWNOKILFEEQDLRK 483
 QY 457 HGLDGEVDSAPLNNIIFQKINCERYYSFHLSEFQEPFAAMYIILDEGEGAGDQDVT 516
 Db 484 HGLDGEVDSAPLNNIIFQKINCERYYSFHLSEFQEPFAAMYIILDEGEGAGDQDVT 543
 QY 517 LLTEYAFSRSFALTGRFLGLNEETRSLEKSLCKWVSPHIMKMDLLOWIQSAQSDG 576
 Db 544 LLTEYAFSRSFALTGRFLGLNEETRSLEKSLCKWVSPHIMKMDLLOWIQSAQSDG 603
 QY 577 STLOQGSLEFSCLEYEQEEFIOQALSHFQVIVVSNIAKWEHWSFCLKRCRSQV 636
 Db 604 STLOQGSLEFSCLEYEQEEFIOQALSHFQVIVVSNIAKWEHWSFCLKRCRSQV 663
 QY 637 HLYGATYSADGEDRARSAGATLLVQLRPTVLLDAYSEHLAAALCTNPNLIELSLYR 696
 Db 664 HLYGATYSADGEDRARSAGATLLVQL-PERTVLLDAYSEHLAAALCTNPNLIELSLYR 722
 QY 697 NALGSRGVKLLCOGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLRTMDLSNGV 756
 Db 723 NALGSRGVKLLCOGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLRTMDLSNGV 782
 QY 757 GFPGMMLLCEGLRHQPCLQMLQRLKQCSGACQEMASVLTGTPHVLVDLITGNAEDL 816
 Db 783 GFPGMMLLCEGLRHQPCLQMLQRLKQCSGACQEMASVLTGTPHVLVDLITGNAEDL 842
 QY 817 GLRLCCQGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLRTMDLSNGV 876
 Db 843 GLRLCCQGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLRTMDLSNGV 902
 QY 877 LLCEGLRHPTCKLQTLRGLICELGSAACEGIVVLOANHNRLDLSFNDLGDGCLWLLA 936
 Db 903 LLCEGLRHPTCKLQTLRGLICELGSAACEGIVVLOANHNRLDLSFNDLGDGCLWLLA 962

QY 937 EGLQHPACRLQKWLDSGGLTAKACENLYFTLGINOTLTDLTLTNALGDTGVRLCKRL 996
 Db 963 EGLQHPACRLQKWLDSGGLTAKACENLYFTLGINOTLTDLTLTNALGDTGVRLCKRL 1022
 QY 997 SHFGCKLRVLWLFQGMIDLNKMTSHSLAALRVTKPYLDIGC 1035
 Db 1023 SHFGCKLRVLWLFQGMIDLNKMTSHSLAALRVTKPYLDIGC 1061
 RESULT 4
 ADC31287
 ID ADC31287 standard; protein; 865 AA.
 XX
 AC ADC31287;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel polypeptide sequence, SEQ ID NO:1369.
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 19.
 XX
 OS Homo sapiens.
 XX WO2003029271-A2.
 XX
 PN 10-APR-2003.
 XX
 PD 24-SEP-2002; 2002WO-US030474.
 XX
 PF 24-SEP-2001; 2001US-0324631P.
 XX
 PR (HYSE-) HYSEQ INC.
 XX
 PA Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC30316.
 PT
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 FS Claim 20; SEQ ID NO 1369; 1185pp; English.
 CC
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the

CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 865 AA;

Query Match 63.9%; Score 661; DB 7; Length 865;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 781; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	113	MEDNRALGECVNLSHRYTRLLLVKHSNPMQVQQQLLTGRGHARTVGHQASPIKIETL	172
Db	1	MEDNRALGECVNLSHRYTRLLLVKHSNPMQVQQQLLTGRGHARTVGHQASPIKIETL	60
QY	173	FPEDBERPPRTVVMQGAAGICKSMIAHKVMDWADGKLFQGRFDYLYINCREMNQSA	232
Db	61	FPEDBERPPRTVVMQGAAGICKSMIAHKVMDWADGKLFQGRFDYLYINCREMNQSA	120
QY	233	TECSMODLIFSCEWPEAPLOELIRVPERLLFIIDGFDLKPSPDQPGPWCWCWEKRP	292
Db	121	TECSMODLIFSCEWPEAPLOELIRVPERLLFIIDGFDLKPSPDQPGPWCWCWEKRP	180
QY	293	TELLNSLRKLLPELSLITTRPTALEKHLRLEHPRHVEITLGFSEAEKKEYFYKYFH	352
Db	181	TELLNSLRKLLPELSLITTRPTALEKHLRLEHPRHVEITLGFSEAEKKEYFYKYFH	240
QY	353	NAEQAGQVNYVRDNEPLFTMCFVPLVWVCTCLOQLEGGGLLQTSRTTAVYMLYL	412
Db	241	NAEQAGQVNYVRDNEPLFTMCFVPLVWVCTCLOQLEGGGLLQTSRTTAVYMLYL	300
QY	413	LSLMQPKGAPRIOPPNEGRLCSLAADGLWQKILFEEQDLRKHGLDGEDVSFAFLNMI	472
Db	301	LSLMQPKGAPRIOPPNEGRLCSLAADGLWQKILFEEQDLRKHGLDGEDVSFAFLNMI	360
QY	473	FKDINCERYYSFIHLSFOEFFAAMYIILDEGGGAGPDQDVTLLTTEYAFSRSFLALT	532
Db	361	FKDINCERYYSFIHLSFOEFFAAMYIILDEGGGAGPDQDVTLLTTEYAFSRSFLALT	420
QY	533	SRFLGLLNEETRSHLEKSLCWKVSPIHKMDLLOWTOSKAQSDGSTLQGSLEFFSCLYE	592
Db	421	SRFLGLLNEETRSHLEKSLCWKVSPIHKMDLLOWTOSKAQSDGSTLQGSLEFFSCLYE	480
QY	593	IQEERFIQALSHFQVIVVSNIAKMEHVMVSSPCLAKRCSAQLVHLGYATSDAGDRAR	652
Db	481	IQEERFIQALSHFQVIVVSNIAKMEHVMVSSPCLAKRCSAQLVHLGYATSDAGDRAR	540
QY	653	CSAGATLLVQLRPERTVLLDAYSEHLAAALCTNPMLIELSLYRNALGSRGVKLLCOGLR	712
Db	541	CSAGATLLVQL-PERTVLLDAYSEHLAAALCTNPMLIELSLYRNALGSRGVKLLCOGLR	599
QY	713	HPNCKLONLRKCRISSSACEDLSAALIANKNLTMDLSGNGVGFPGMWMLCEGRHPQ	772
Db	600	HPNCKLONLRKCRISSSACEDLSAALIANKNLTMDLSGNGVGFPGMWMLCEGRHPQ	659
QY	773	CRLOMTQLRKQLESACQEMASVLTGNPHLVELDTGNALEDGLRLLCQGLRHVPVRL	832
Db	660	CRLOMTQLRKQLESACQEMASVLTGNPHLVELDTGNALEDGLRLLCQGLRHVPVRL	719
QY	833	RTLWLKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGRHPTCKLQTL	892
Db	720	RTLWLKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGRHPTCKLQTL	779

QY	893	RL 894
Db	780	RL 781
RESULT 5		
ID	ADE36457	standard; protein; 603 AA.
XX	ADE36457;	
AC	ADE36457;	
XX	29-JAN-2004	(first entry)
DT		
XX	Human PAN6 leucine-rich-repeat domain (LRR) seq id 64.	
DE		
XX		
XX		
KW	cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;	
KW	antiallergic; antitumor; dermatological; cerebroprotective; cardiact;	
KW	antiParkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;	
KW	NFKappaB activation inhibitor; PAAD domain containing polypeptide;	
KW	PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;	
KW	apoptosis-associated speck-like protein; caspase recruitment domain 2;	
KW	ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;	
KW	apoptosis; NFKappaB induction; cytokine processing;	
KW	cytokine receptor signaling caspase-mediated proteolysis;	
KW	c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;	
KW	inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;	
KW	neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;	
KW	fibrosis; smooth muscle cell proliferation; balloon angioplasty;	
KW	restenosis; leukaemia; lymphoma; inflammatory disease; allergy;	
KW	arthritis; lupus; schroen's syndrome; Crohn's disease;	
KW	ulcerative colitis; graft versus host disease; stroke; heart failure;	
KW	neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;	
KW	cancer therapy; PAAD domain family; human; PAN6; leucine-rich-repeat;	
KW	LLR.	
XX		
OS	Homo sapiens.	
XX		
PN	US2003077699-A1.	
XX		
PD	24-APR-2003.	
XX		
PF	25-SEP-2001; 2001US-00965621.	
XX		
PR	26-SEP-2000; 2000US-00671760.	
PR	26-SEP-2000; 2000US-0367367E.	
XX		
PA	(REED/) REED J C.	
PA	(GODZ/) GODZIK A.	
PA	(CHUZ/) CHU Z.	
PA	(PAWL/) PAWLÓWSKI K.	
PA	(FIOR/) FIORENTINO L.	
PA	(ARIZ/) ARIZA M E.	
PA	(STEH/) STEHLIK C.	
XX		
PI	Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;	
PI	Stehlik C;	
XX		
DR	WPI; 2002-471256/50.	
XX		
PT	Novel isolated PAAD domain containing polypeptide useful for inducing	
PT	apoptosis by inhibiting nuclear factor kappa B activation and in gene	
PT	therapy for treating cancer.	
XX		
PS	Claim 24; SEQ ID NO 64; 93pp; English.	
XX		
CC	The invention describes an isolated PAAD domain containing polypeptide	
CC	(I) comprising 80% identity to the amino acid sequence of PAAD and	
CC	nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck	
CC	-like protein containing a caspase recruitment domain (ASC)-2 fully	
CC	defined in specification, where (I) is biologically active. (I) is useful	
CC	for identifying a (I)-associated polypeptide, an agent altering that	
CC	association and agents that modulate PAAD domain mediated inhibition of	
CC	nuclear factor kappa B (NFKappaB). A NB-ARC domain polypeptide is useful	

CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (1). (1) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
 CC processing, cytokine receptor signaling, caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (1) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (1) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6 leucine-rich-repeat domain (LLR).
 XX
 SQ Sequence 603 AA;

Query Match 58.3%; Score 603; DB 5; Length 603;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNORGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNMNIFQDINCERYYSIHL 488
 Db 1 PNORGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNMNIFQDINCERYYSIHL 60
 QY 489 SPOEFFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFALTSRFLGLNBEETRSHL 548
 Db 61 SPOEFFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFALTSRFLGLNBEETRSHL 120
 QY 549 EKSICWKVSPHIXMDLLOWTOSKAQSDGSTLOQSLIEFFSCLVEIOEEFTQOALSHPQV 608
 Db 121 EKSICWKVSPHIXMDLLOWTOSKAQSDGSTLOQSLIEFFSCLVEIOEEFTQOALSHPQV 180
 QY 609 IVVSNIAKMEHVMVSSCLKRCRQAQVHLHYGATYSADGEDRARCAGAHNTLLVQLRPER 668
 Db 181 IVVSNIAKMEHVMVSSCLKRCRQAQVHLHYGATYSADGEDRARCAGAHNTLLVQLRPER 240
 QY 669 TVLLDAYSEHLAALCTNPMLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRI 728
 Db 241 TVLLDAYSEHLAALCTNPMLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRI 300
 QY 729 SSSACEDLSAALTANKNLTRMDLSGNGVGPWMLLCEGLRHPNCKLQNLRLKRCRI 788
 Db 301 SSSACEDLSAALTANKNLTRMDLSGNGVGPWMLLCEGLRHPNCKLQNLRLKRCRI 360
 QY 789 ACQEMASVLGTNPMLVELDTGNALEDGLRLCQGLRHPNCKLQNLRLKRCRI 848
 Db 361 ACQEMASVLGTNPMLVELDTGNALEDGLRLCQGLRHPNCKLQNLRLKRCRI 420
 QY 849 ELASTLSVNSQSLRELDLSINELGDLGVLLCEGLRHPNCKLQNLRLKRCRI 908
 Db 421 ELASTLSVNSQSLRELDLSINELGDLGVLLCEGLRHPNCKLQNLRLKRCRI 480
 QY 909 VVLQANHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKMLDSCGLTAKACENLYFTL 968
 Db 481 VVLQANHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKMLDSCGLTAKACENLYFTL 540
 QY 969 GINQTLTDLVLTNNALGDTGVRLLCRSLSHPGCKLRVLWLFQMDLNKMTSRLAALRVTK 1028
 Db 541 GINQTLTDLVLTNNALGDTGVRLLCRSLSHPGCKLRVLWLFQMDLNKMTSRLAALRVTK 600
 QY 1029 PYL 1031
 Db 601 PYL 603

RESULT 6
 ABG97475
 ID ABG97475 standard; protein; 565 AA.
 XX

AC ABG97475;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human nucleic acid associated protein, NAAP9, from, INCYTE no.429930CD1.
 XX
 KW Human; nucleic acid associated protein; NAAP; cancer;
 KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
 KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
 KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
 KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
 KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
 KW atopic dermatitis; arthritis; bacterial infection; viral infection;
 KW parasitic infection; protozoal infection; fungal infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200272630-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 07-FEB-2002; 2002WO-US003844.
 XX
 PR 09-FEB-2001; 2001US-0268118P.
 PR 21-FEB-2001; 2001US-0270963P.
 PR 22-FEB-2001; 2001US-0270858P.
 PR 23-FEB-2001; 2001US-0271194P.
 PR 07-MAR-2001; 2001US-0274071P.
 PR 12-APR-2001; 2001US-0283496P.
 PR 09-NOV-2001; 2001US-0344650P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
 PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
 PI Gandhi AR, Ding L, Yue H, Gietzen KU, Wallia NK, Thangavelu K;
 PI Elliott VS, Marquis JP;
 XX
 DR WPI: 2002-723320/78.
 DR N-PSDB; ABS78719.
 XX
 FT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
 FT treating and preventing diseases or conditions associated with the
 FT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
 FT infections.
 XX
 PS Claim 1; Page 147-149; 162pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising one of 10
 CC human nucleic acid associated protein (NAAP1-10), or a biologically
 CC active or immunogenic fragment of the polypeptide, and their encoding
 CC nucleic acid. Also included are a recombinant polynucleotide comprising a
 CC promoter sequence operably linked to the polynucleotide, a cell
 CC transformed with the recombinant polynucleotide, a transgenic organism
 CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
 CC screening for a compound that is effective as an ant/agonist or modulator
 CC of NAAP, generating an expression profile of a sample containing the
 CC polynucleotides and an array comprising different nucleotide molecules
 CC affixed on a solid substrate, nucleotide molecule comprises a first
 CC oligonucleotide or polynucleotide sequence specifically hybridisable with
 CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
 CC The polypeptides and polynucleotides are useful in diagnosing, treating
 CC and preventing diseases or conditions associated with the decreased
 CC expression or overexpression of NAAP, such as cell proliferative diseases
 CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
 CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
 CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
 CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
 CC other diseases and disorders listed in the specification. These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its

CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence represents an NAAP protein
 XX
 SQ Sequence 565 AA;

Query Match 43.0%; Score 445; DB 5; Length 565;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 470 MNIFQKINCERYYSFIHLSFQEFFRMYVILDEGGGAGPDQVTRLLTEYAFSRSFL 529
 Db 1 MNIFQKINCERYYSFIHLSFQEFFRMYVILDEGGGAGPDQVTRLLTEYAFSRSFL 60
 QY 530 ALTSRFLGGLNEETSHLEKSLCWKVSPIHKNMQLQWIOSKAQSDGSTLQOGLSFFSC 589
 Db 61 ALTSRFLGGLNEETSHLEKSLCWKVSPIHKNMQLQWIOSKAQSDGSTLQOGLSFFSC 120
 QY 590 LYEIOBEETIQALSHFQVIVSNIAKMEHMYVSFCLKRCRQAQVHLHYGATYSADGED 649
 Db 121 LYEIOBEETIQALSHFQVIVSNIAKMEHMYVSFCLKRCRQAQVHLHYGATYSADGED 180
 QY 650 RARCSAGATLLVQLPERTVLLDAYSEHLAALCTNPNIJELSLYRNALGSGVKLLCQ 709
 Db 181 RARCSAGATLLVQLPERTVLLDAYSEHLAALCTNPNIJELSLYRNALGSGVKLLCQ 239
 QY 710 GLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 769
 Db 240 GLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 299
 QY 770 HPQCRLOMTQLKRCQLESAGCQEMASVLTGNPHLVELDTGNALDGLRLCQGLRHPV 829
 Db 300 HPQCRLOMTQLKRCQLESAGCQEMASVLTGNPHLVELDTGNALDGLRLCQGLRHPV 359
 QY 830 CRILTLMWKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGLRHPTCKL 889
 Db 360 CRILTLMWKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGLRHPTCKL 419
 QY 890 QTLRLGICRLGSAACBGLSVLQANLNRELDLSFNDLGDWGLWLLAEGLOHPACRLQKL 949
 Db 420 QTLRLGICRLGSAACBGLSVLQANLNRELDLSFNDLGDWGLWLLAEGLOHPACRLQKL 479
 QY 950 WLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRSLHPGCKLRVLWLF 1009
 Db 480 WLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRSLHPGCKLRVLWLF 539
 QY 1010 GMDLNKMTSHRLAALRVTKPYLDIGC 1035
 Db 540 GMDLNKMTSHRLAALRVTKPYLDIGC 565

RESULT 7
 ID ABU99120 standard; protein; 521 AA.
 XX AC ABU99120;
 XX DT 01-AUG-2003 (first entry)
 XX DE Novel human GPCR related protein NOV2b.
 XX KW Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytosolic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antilasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;

KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX Homo sapiens.
 PN WO200299116-A2.
 XX 12-DEC-2002.
 XX 04-JUN-2002; 2002WO-US017428.
 PF 04-JUN-2001; 2001US-0295607P.
 XX 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.
 PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.
 XX (CURA-) CURAGEN CORP.
 XX Anderson DW, Baumgartner JC, Boldog FL, Casman SU, Edinger SR;
 PI Ganggalli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;
 PI Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 PI Voss EZ, Zerhusen BD;
 XX WPI; 2003-140627/13.
 DR N-PSDB; ACD03624.
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 1; Page 100; 332pp; English.
 XX The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods. The nucleic acids are further used as
 CC hybridisation probes in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The polypeptides are also useful as
 CC vaccines. This is the amino acid sequence of a novel human G-protein
 CC coupled receptor related protein NOV
 XX Sequence 521 AA;

Query Match	39.4%;	Score 408;	DB 6;	Length 521;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches	408;	Conservative	0;	Mismatches	0;
QY	373	MCVPLVWVCTCLOQLEGGGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQR	432		
DB	1	MCVPLVWVCTCLOQLEGGGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQR	60		
QY	433	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE	492		
DB	61	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE	120		
QY	493	FFAAMYILDEGGGAGPDQDVTLTTEYAFSERSFLATSRFLGFLNNEETSHLEKSL	552		
DB	121	FFAAMYILDEGGGAGPDQDVTLTTEYAFSERSFLATSRFLGFLNNEETSHLEKSL	180		
QY	553	CKVSPHIKMDLLQWIOSKAQSDGSTLQOQSLEFFSCLYEIQEERTIQOALSHPQIVVVS	612		
DB	181	CKVSPHIKMDLLQWIOSKAQSDGSTLQOQSLEFFSCLYEIQEERTIQOALSHPQIVVVS	240		
QY	613	NIASKMEHVMVSFCLKRCRAQVLLHYGATYSADGEDRARCAGAHLLVQLRPTVLL	672		
DB	241	NIASKMEHVMVSFCLKRCRAQVLLHYGATYSADGEDRARCAGAHLLVQLRPTVLL	300		
QY	673	DAYSEHLAALCTNPNIIELSLRNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSA	732		
DB	301	DAYSEHLAALCTNPNIIELSLRNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSA	360		
QY	733	CEDLSAALIANKNLTRMDLSGNGVGFPCWMLLCEGLRHPQCRLOMIQL	780		
DB	361	CEDLSAALIANKNLTRMDLSGNGVGFPCWMLLCEGLRHPQCRLOMIQL	408		
RESULT 8					
ID	ABU99119	standard; protein; 582 AA.			
XX	AC	ABU99119;			
XX	AC	ABU99119;			
DT	01-AUG-2003	(first entry)			
XX	DE	Novel human GPCR related protein NOV2a.			
XX	KW	Human, G-protein coupled receptor related protein; GPCR related protein;			
KW	NOV;	cytostatic; cardiant; antiarteriosclerotic; antidiabetic;			
KW	immunomodulator;	anti-HIV; anorectic; antiasthmatic; haemostatic;			
KW	antiparkinsonian;	neuroprotective; neurotropic; gene therapy; vaccine;			
KW	NOVX-associated disorder;	cardiomyopathy; atherosclerosis; cancer;			
KW	diabetes;	immune disorder; AIDS; obesity; asthma;			
KW	haematopoietic disorder;	Parkinson's disease; Alzheimer's disease;			
KW	infection;	multiple sclerosis; cancer-associated cachexia;			
KW	wasting disorder;	chronic disease; neurogenesis; cell differentiation;			
KW	cell proliferation;	haematopoiesis; wound healing; angiogenesis;			
KW	chromosome mapping;	tissue typing; preventive medicine; pharmacogenomic.			
OS	Homo sapiens.				
XX	XX				
PN	WO200299116-A2.				
XX	XX				
PD	12-DEC-2002.				
XX	XX				
PF	04-JUN-2002;	2002WO-US017428.			
XX	XX				
PR	04-JUN-2001;	2001US-0295607P.			
PR	04-JUN-2001;	2001US-0295661P.			
PR	06-JUN-2001;	2001US-0296404P.			
PR	06-JUN-2001;	2001US-0296418P.			
PR	14-JUN-2001;	2001US-0298285P.			
PR	15-JUN-2001;	2001US-0298556P.			
PR	21-JUN-2001;	2001US-0299949P.			
PR	26-JUN-2001;	2001US-0300883P.			
PR	28-JUN-2001;	2001US-0301550P.			
PR	13-AUG-2001;	2001US-03111972P.			
PR	27-AUG-2001;	2001US-0315071P.			
PR	29-AUG-2001;	2001US-0315660P.			
PR	14-SEP-2001;	2001US-0322293P.			
PR	17-SEP-2001;	2001US-0322706P.			
PR	14-DEC-2001;	2001US-0341186P.			
PR	28-FEB-2002;	2002US-0361189P.			
PR	12-MAR-2002;	2002US-0363673P.			
PR	12-MAR-2002;	2002US-0363876P.			
PR	03-JUN-2002;	2002US-00363676.			
XX		(CURA-) CURAGEN CORP.			
XX					
PI	Anderson DW,	Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;			
PI	Gangolli EA,	Gerlach VL, Gorman L, Guo X, Hjalte T, Kekuda R, Li L;			
PI	Maddougali JR,	Malyankar UM, Millet I, Padigaru M, Patturajan M;			
PI	Pena CE,	Rastelli L, Shinkets RA, Stone DU, Spytek KA, Vernet CAM;			
PI	Voss EZ,	Zerhusen BD;			
XX					
DR	WPI;	2003-140627/13.			
DR	N-PSDB;	ACD03623.			
XX					
XX		New NOVX polypeptides and nucleic acids, useful for preventing or			
PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,				
PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or				
PT	pharmacogenomics.				
XX					
PS	Claim 1;	Page 99; 332pp; English.			
XX					
CC		The invention describes an isolated polypeptide (I) comprising any of 27			
CC		118-961 residue amino acid sequences, given in the specification, a			
CC		mature form of them, a sequence that is at least 95 % identical to them,			
CC		or a sequence having one or more conservative substitutions in them. The			
CC		polypeptide is useful in manufacturing a medicament for treating a			
CC		polysome associated with a human disease selected from a pathology			
CC		associated with the polypeptide. The NOVX polypeptides, polynucleotides			
CC		and antibodies are useful in treating or preventing NOVX-associated			
CC		disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune			
CC		diseases, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's			
CC		disease, Alzheimer's disease, infections, multiple sclerosis, cancer-			
CC		associated cachexia, and other wasting disorders associated with chronic			
CC		diseases. The nucleic acids and polypeptides may also be used as targets			
CC		for the identification of small molecules that modulate or inhibit e.g.			
CC		neurogenesis, cell differentiation, cell proliferation, haematopoiesis,			
CC		wound healing and angiogenesis, in gene therapy, in generation of			
CC		antibodies that bind immunospecifically to NOVX substances for use in			
CC		therapeutic or diagnostic methods. The nucleic acids are further used as			
CC		hybridisation probes, in chromosome mapping, tissue typing, preventive			
CC		medicine, and pharmacogenomics. The polypeptides are also useful as			
CC		vaccines. This is the amino acid sequence of a novel human G-protein			
CC		coupled receptor related protein NOV			
XX					
SQ		Sequence 582 AA;			
Query Match	33.8%;	Score 350;	DB 6;	Length 582;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches	350;	Conservative	0;	Mismatches	0;
QY	373	MCVPLVWVCTCLOQLEGGGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQR	432		
DB	1	MCVPLVWVCTCLOQLEGGGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQR	60		
QY	433	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE	492		
DB	61	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE	120		
QY	493	FFAAMYILDEGGGAGPDQDVTLTTEYAFSERSFLATSRFLGFLNNEETSHLEKSL	552		
DB	121	FFAAMYILDEGGGAGPDQDVTLTTEYAFSERSFLATSRFLGFLNNEETSHLEKSL	180		
QY	553	CKVSPHIKMDLLQWIOSKAQSDGSTLQOQSLEFFSCLYEIQEERTIQOALSHPQIVVVS	612		
DB	181	CKVSPHIKMDLLQWIOSKAQSDGSTLQOQSLEFFSCLYEIQEERTIQOALSHPQIVVVS	240		

QY 613 NTASKMEHVSFCLRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPRTVLL 672
 Db 241 NTASKMEHVSFCLRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPRTVLL 300
 QY 673 DAYSEHLAALCTNPNIELSLRYNALGSRGVKLLCQGLRHFNCKLQNL 722
 Db 301 DAYSEHLAALCTNPNIELSLRYNALGSRGVKLLCQGLRHFNCKLQNL 350

RESULT 9
 ABB77910
 ID ABB77910 standard; protein; 449 AA.
 XX AC ABB77910;
 XX DT 07-OCT-2002 (first entry)
 XX Amino acid sequence of human leucine-rich repeat protein HLRBEM1.
 DE Human; leucine-rich repeat; HLRBEM1; proliferative disorder;
 XX Human condition; apoptosis; signal transduction; autoimmune disease;
 XX haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
 KW cardiovascular disease; neurological disease; pheromone;
 KW pulmonary disease; chronic obstructive pulmonary disease;
 KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
 KW haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
 KW inflammatory disorder; systemic lupus erythematosus;
 KW cardiovascular disease; cancer.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Domain 144..159
 FT /note= "transmembrane domain"
 FT
 PN WO200252011-A2.
 XX
 XX 04-JUL-2002.
 XX
 XX 20-DEC-2001; 2001WO-US049740.
 XX
 XX 22-DEC-2000; 2000US-0257773P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Feder J, Ramanathan C, Mintier G;
 PI
 XX WPI; 2002-566676/60.
 DR N-PSDB; ABI59333.
 XX
 XX New HLRBEM1 nucleic acids for preventing, treating or ameliorating e.g.
 FT proliferative disorders, immune conditions, a disorder related to
 FT aberrant apoptosis modulation or developmental disorders.
 XX
 XX Claim 20; Fig 1A-E; 371pp; English.
 PS
 XX The present sequence represents a human leucine-rich repeat containing
 CC protein, designated HLRBEM1. HLRBEM1 polypeptides and polynucleotides are
 CC useful for preventing, treating or ameliorating a medical condition such
 CC as a proliferative disorder, immune condition, or a disorder related to
 CC aberrant apoptosis modulation, either directly or indirectly, and in
 CC modulating signal transduction activity in various cells, tissue and
 CC organisms. They are also useful for treating, preventing, or diagnosing
 CC diseases of haematopoietic cells, autoimmune disease, graft-versus-host
 CC disease, allergic conditions (e.g. asthma), cardiovascular disorders, and
 CC neurological diseases, and for increasing the organisms' ability to
 CC synthesize and/or release pheromones. The polypeptide may also be used in
 CC treating, preventing or ameliorating pulmonary disease (e.g. chronic
 CC obstructive pulmonary disease, allergic rhinitis, or bronchial
 CC hyperresponsiveness), reproductive disease, haematopoietic disease,
 CC platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious
 CC disorders (e.g. innate immunity to bacterial pathogens, or adaptive

CC immune response), immune and inflammatory disorders (e.g. systemic lupus
 CC erythematosus), cardiovascular diseases and cancers. HLRBEM1 nucleic
 CC acids may further be used in chromosome identification or mapping, as a
 CC chromosome marker, as molecular weight markers, as diagnostic probes, in
 CC gene therapy, in raising anti-DNA antibodies, or as antigens for
 CC eliciting immune responses
 XX
 SQ Sequence 449 AA;

Query Match 24.5%; Score 254; DB 5; Length 449;
 Best Local Similarity 100.0%; Pred. No. 2.1e-243;
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 MNQSAATECSMDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGPWCLW 287
 Db 1 MNQSAATECSMDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGPWCLW 60
 QY 288 EEKRPTELLNSLRKKLPPELSLLITTRPTALEKHLRLLEHPRHVEILGFSAEERKEYF 347
 Db 61 EEKRPTELLNSLRKKLPPELSLLITTRPTALEKHLRLLEHPRHVEILGFSAEERKEYF 120
 QY 348 KYFHNAAEQAGQVFNVRDNEPLFTMCVFPLVCVVVCTCQQQLEGGLRQTSRTTAV 407
 Db 121 KYFHNAAEQAGQVFNVRDNEPLFTMCVFPLVCVVVCTCQQQLEGGLRQTSRTTAV 180
 QY 408 YMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHGLDGEDVSFAF 467
 Db 181 YMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHGLDGEDVSFAF 240
 QY 468 LNMNIFQKINCER 481
 Db 241 LNMNIFQKINCER 254

RESULT 10

AAE21064
 ID AAE21064 standard; protein; 287 AA.

XX AC AAE21064;

XX DT 01-JUL-2002 (first entry)

XX rno (upregulated by nitric oxide)-3 protein.

XX Upregulated by nitric oxide; NO; rno-3; differentiation; death; cancer;
 XX acute myeloid leukaemia; AML; cytostatic.

XX OS Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 249

FT /label= Unknown
 FT /note= "This residue is represented in the specification as O"

FT Misc-difference 256

FT /label= Unknown
 FT /note= "This residue is represented in the specification as O"

XX US2001029033-A1.

XX 11-OCT-2001.

XX 06-MAR-2001; 2001US-00799983.

XX 06-MAR-2000; 2000US-0186971P.

XX (SHAM/) SHAMI P J.

XX (PARK/) PARKER C J.

XX Shami PJ, Parker CJ;

XX WPI; 2002-009982/01.

XX Novel polypeptide regulated by nitric oxide useful for inducing acute
PT myeloid leukemia cell differentiation and apoptosis.
XX
XX
XX Claim 29; Page; 23pp; English.
XX
XX The present invention relates to novel genes which are upregulated by
CC nitric oxide (NO), designated as rno and their corresponding proteins.
CC The invention also relates to the isolation and characterisation of three
CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
CC are useful for inducing differentiation and death in a cancer cell.
CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
CC expression of a rho gene comprising the nucleotides coding for rno
CC protein. The present sequence is rno-3 protein. Note: This sequence is
CC not shown in the specification but is derived from rno-1 protein shown as
CC SRQ ID NO: 2 in figure 7B of the specification
XX
XX
XX Sequence 287 AA;

Query Match 21.0%; Score 217; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 9.5e-207; Mismatches 0; Indels 0; Gaps 0;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 723 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIOLRK 782
Db 32 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIOLRK 91
QY 783 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 842
Db 92 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 151
QY 843 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTOTLRLGICRLGSA 902
Db 152 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTOTLRLGICRLGSA 211
QY 903 ACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEG 939
Db 212 ACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEG 248

RESULT 11
AAE21062
ID AAE21062 standard; protein; 344 AA.
XX
XX AAE21062;
XX
XX 01-JUL-2002 (first entry)
XX
XX rno (upregulated by nitric oxide)-1 protein.
XX
XX Upregulated by nitric oxide; NO; rno-1; differentiation; death; cancer;
KW acute myeloid leukaemia; AML; cytostatic.
XX
XX Unidentified.

Key Location/Qualifiers
FT Misc-difference 249
FT /label= Unknown
FT /note= "This residue is represented in the specification
FT as O"
FT Misc-difference 256
FT /label= Unknown
FT /note= "This residue is represented in the specification
FT as O"
XX
XX US2001029033-A1.
XX
XX 11-OCT-2001.
XX
XX 06-MAR-2001; 2001US-00799983.
XX
XX 06-MAR-2000; 2000US-0186971P.
XX

PA (SHAM/) SHAMI P J.
XX (PARK/) PARKER C J.
XX
XX Shami PU, Parker CJ;
XX
XX MPI; 2002-009982/01.
XX
XX Novel polypeptide regulated by nitric oxide useful for inducing acute
PT myeloid leukemia cell differentiation and apoptosis.
XX
XX Claim 29; Fig 7B; 23pp; English.
XX
XX The present invention relates to novel genes which are upregulated by
CC nitric oxide (NO), designated as rno and their corresponding proteins.
CC The invention also relates to the isolation and characterisation of three
CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
CC are useful for inducing differentiation and death in a cancer cell.
CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
CC expression of a rho gene comprising the nucleotides coding for rno
CC protein. The present sequence is rno-1 protein
XX
XX Sequence 344 AA;

Query Match 21.0%; Score 217; DB 5; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.1e-206;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 723 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIOLRK 782
Db 32 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIOLRK 91
QY 783 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 842
Db 92 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 151
QY 843 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTOTLRLGICRLGSA 902
Db 152 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTOTLRLGICRLGSA 211
QY 903 ACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEG 939
Db 212 ACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEG 248

RESULT 12
ADE36456
ID ADE36456 standard; protein; 190 AA.
XX
XX ADE36456;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human PAN6 NB-ARC domain seq id 63.

XX
KW cytostatic; immunosuppressive; vulnary; antinflammatory; vasotropic;
KW anti-allergic; antiulcer; dermatological; cerebroprotective; cardiac;
KW antiparkinsonian; neurotropic; neuroprotective; anti-HIV; gene therapy;
KW NPKappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NPKappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; PAN6; NB-ARC domain.
XX

OS Homo sapiens.
 XX US2003077699-A1.
 PN XX
 PD XX
 PD 24-APR-2003.
 XX
 XX 25-SEP-2001; 2001US-00965621.
 XX
 XX 26-SEP-2000; 2000US-00671760.
 PR 26-SEP-2000; 2000US-0367367F.
 XX
 XX (REED/) REED J C.
 PA (GODZ/) GODZIK A.
 PA (CHUZ/) CHU Z.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (ARIZ/) ARIZA M E.
 PA (STEH/) STEHLIK C.
 XX
 XX Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;
 XX
 XX WPI; 2002-471256/50.
 DR
 XX Novel isolated PAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 XX Claim 22; SEQ ID NO 63; 93pp; English.
 XX
 XX The invention describes an isolated PAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAD domain mediated inhibition of
 CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical processes such as apoptosis, NFkappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAD and nucleotide
 CC binding protein PAN6 NB-ARC domain which demonstrates nucleotide binding
 CC and facilitates the binding to another NB-ARC containing protein.
 XX
 XX Sequence 190 AA;
 SQ
 Query Match 18.4%; Score 190; DB 5; Length 190;
 Best Local Similarity 100.0%; Pred. No. 4.9e-180;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 183 PRTVMQGAAGIGKSLAHKWLMDWADGKLFQGRFDYLFYINCREMNQATSCSMQDLIF 242
 DB 1 PRTVMQGAAGIGKSLAHKWLMDWADGKLFQGRFDYLFYINCREMNQATSCSMQDLIF 60
 243 SCWPESAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLCWEKPEPTLLNSLIR 302
 DB 61 SCWPESAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLCWEKPEPTLLNSLIR 120
 303 KKLLPELSLLITRPTALEKHLRHHVILGFSEAEKKEYFYKFFHNAEQAGQVFN 362

DB 121 KKLLPELSLLITRPTALEKHLRHHVILGFSEAEKKEYFYKFFHNAEQAGQVFN 180
 QY 363 YVRDNEPLFT 372
 DB 181 YVRDNEPLFT 190
 RESULT 13
 ADC32968
 ID ADC32968 standard; protein; 197 AA.
 XX
 AC ADC32968;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3050.
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 19.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC32201.
 XX
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Example 2; SEQ ID NO 3050; 1185pp; English.
 XX
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; a method of detecting
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are

CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 197 AA;

Query Match 18.4%; Score 190; DB 7; Length 197;
 Best Local Similarity 100.0%; Pred. No. 5.1e-180;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 693 SLRYNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 752
 Db 1 SLRYNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 60
 QY 753 GNGVGFPMMLLCEGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 812
 Db 61 GNGVGFPMMLLCEGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 120
 QY 813 LEDGLRLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 872
 Db 121 LEDGLRLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 180
 QY 873 LGVLLLCCEGL 882
 Db 181 LGVLLLCCEGL 190

RESULT 14
 AAE21063
 ID AAE21063 standard; protein; 287 AA.
 XX
 AC AAE21063;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE rno (upregulated by nitric oxide)-2 protein.
 XX
 KW Upregulated by nitric oxide; NO; rno-2; differentiation; death; cancer;
 KW acute myeloid leukaemia; AML; cytostatic.
 XX
 OS Unidentified.
 XX
 PN US2001029033-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 06-MAR-2001; 2001US-00799983.
 XX
 PR 06-MAR-2000; 2000US-0186971P.
 XX
 PA (SHAM)/ SHAMI P J.
 PA (PARK)/ PARKER C J.
 XX
 PI Shami PJ, Parker CJ;
 XX
 DR WPI; 2002-009982/01.
 XX
 CC Novel polypeptide regulated by nitric oxide useful for inducing acute
 CC myeloid leukemia cell differentiation and apoptosis.
 XX
 PS Claim 29; Page; 23pp; English.
 XX
 CC The present invention relates to novel genes which are upregulated by
 CC nitric oxide (NO), designated as rno and their corresponding proteins.
 CC The invention also relates to the isolation and characterisation of three

CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
 CC are useful for inducing differentiation and death in a cancer cell.
 CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
 CC expression of a rno gene comprising the nucleotides coding for rno
 CC protein. The present sequence is rno-2 protein. Note: This sequence is
 CC not shown in the specification but is derived from rno-1 protein shown as
 CC SEQ ID NO: 2 in figure 7B of the specification
 XX
 SQ Sequence 287 AA;

Query Match 16.6%; Score 172; DB 5; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.9e-162;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 723 LKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPMMLLCEGLRHPNCKLQNLK 782
 Db 32 LKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPMMLLCEGLRHPNCKLQNLK 91
 QY 783 CQLESGACQEMASVLTGTPHVLVDLTGNALEDGLRLCQGLRHPNCKLQNLK 842
 Db 92 CQLESGACQEMASVLTGTPHVLVDLTGNALEDGLRLCQGLRHPNCKLQNLK 151
 QY 843 TAAACDELASTLSVNQSLRELDLSNELDGLVLLLCCEGLRHPNCKLQNLK 894
 Db 152 TAAACDELASTLSVNQSLRELDLSNELDGLVLLLCCEGLRHPNCKLQNLK 203

RESULT 15
 ABG04570
 ID ABG04570 standard; protein; 168 AA.
 XX
 AC ABG04570;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #4561.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS68757.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 34929; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 168 AA;

Query Match 12.2%; Score 126; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 2.2e-116;
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 539 LNEETRSHLEKSLCWKSPHIXMDLLOWIQSKAQSDGSTLQQGSLEFFSCLYEIOEEF 598
 Db 12 LNEETRSHLEKSLCWKSPHIXMDLLOWIQSKAQSDGSTLQQGSLEFFSCLYEIOEEF 71
 QY 599 IQQALSHFQVIVVSNIAKMEHNVSSFCIKRCSAQVLHLYGATYSADGEDRARCAGAH 658
 Db 72 IQQALSHFQVIVVSNIAKMEHNVSSFCIKRCSAQVLHLYGATYSADGEDRARCAGAH 131
 QY 659 TLIVQL 664
 Db 132 TLIVQL 137

Search completed: July 30, 2004, 14:02:48
 Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 14:00:43 ; Search time 21 Seconds
(without alignments)
4740.868 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 1035
Sequence: 1 MLRTAGDGLCRSLTYLEEL.....MTHSRALALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.9	1192	T17255	hypothetical prote
2	8	0.8	87	AC0434	Bola-like protein
3	8	0.8	122	S60262	corticotensin - ra
4	8	0.8	123	D71922	probable flagellar
5	8	0.8	123	H64592	flagellar switch p
6	8	0.8	313	B95264	probable LysR-fam
7	8	0.8	360	S68209	ads22 protein homo
8	8	0.8	456	S20597	ribonuclease inhib
9	8	0.8	456	A31857	ribonuclease inhib
10	8	0.8	461	A31858	ribonuclease angio
11	8	0.8	475	Z4BPT4	helicase (EC 3.6.1
12	8	0.8	492	F86263	hypothetical prote
13	8	0.8	512	T38422	probable chromatin
14	8	0.8	611	A11941	hypothetical prote
15	8	0.8	661	I56258	RP105 - mouse
16	8	0.8	978	E96787	protein T4012.5 [i
17	8	0.8	1784	C96615	hypothetical prote
18	7	0.7	36	B44400	myosin heavy chain
19	7	0.7	50	B82279	hypothetical prote
20	7	0.7	54	S31533	ribulose-bisphosph
21	7	0.7	89	H82166	hypothetical prote
22	7	0.7	100	D41608	hima protein - Rho
23	7	0.7	100	T49857	hypothetical prote
24	7	0.7	102	W4WL31	E4 protein - human
25	7	0.7	116	S62513	hypothetical 13.8K
26	7	0.7	121	A84083	hypothetical prote
27	7	0.7	135	F97942	hypothetical prote
28	7	0.7	141	F75252	hypothetical prote
29	7	0.7	142	AB1635	hypothetical prote

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFP58601822.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17255
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: EMBL:AL117470
A:Experimental source: adult uterus; clone DKFP58601822
C:Genetics:
A>Note: DKFP58601822.1

Query Match 0.9%; Score 9; DB 2; Length 1192;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 QGAAGIGKS 197
DB 52 QGAAGIGKS 60

RESULT 2

AC0434
Bola-like protein [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0434
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92799.1; PID:G15981492; GSPDB:GN00175
C:Genetics:
A:Gene: YPO3570
C:Superfamily: bola protein

Query Match 0.8%; Score 8; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

30 7 0.7 154 2 C81579 conserved hypotet
31 7 0.7 159 2 H75391 hypothetcal prote
32 7 0.7 171 2 D72090 hypothetcal prote
33 7 0.7 171 2 F86532 hypothetcal prote
34 7 0.7 179 2 F75519 Holliday junction
35 7 0.7 192 2 T06284 hypothetcal prote
36 7 0.7 194 2 JC5260 progesterone membr
37 7 0.7 196 2 AI0555 4-methyl-5(b-hydro
38 7 0.7 197 2 H86741 conserved hypotet
39 7 0.7 198 1 H64771 hydroxymethylpyrim
40 7 0.7 198 2 F90688 hypothetcal prote
41 7 0.7 198 2 B85539 hypothetcal prote
42 7 0.7 208 2 AE2806 glutathione S-tran
43 7 0.7 208 2 D97585 probable DNA-3-met
44 7 0.7 215 2 G98285 DNA-3-methyladenin
45 7 0.7 218 2 AB2998

QY 604 SHFQIVV 611
 Db 29 SHFQIVV 36

RESULT 3

S60262
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jan-2003
 C:Accession: S60262
 R:Vaughan, J.; Donaldson, C.; Bittencourt, J.; Perrin, M.H.; Lewis, K.; Sutton, S.; Chan
 Nature 378, 287-292, 1995
 A:Title: Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticot
 A:Reference number: S60262; MUID:96069764; PMID:7477349
 A:Accession: S60262
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-122 <VAN>
 A:Cross-references: EMBL:U33935; NID:G1065908; PIDN:AAA87566.1; PID:G1065909
 C:Superfamily: corticotiberin-endorpholiberin; diuretic hormone homology
 F;80-120/Domain: diuretic hormone homology <DHH>

Query Match 0.8%; Score 8; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 LLVQLRPE 667
 Db 15 LLVQLRPE 22

RESULT 4

D71922
 probable flagellar motor switch protein - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: Strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
 C:Accession: D71922
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: D71922
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <ARN>
 A:Cross-references: GB:AE001485; GB:AE001439; NID:G4155057; PIDN:AAD06097.1; PID:G415506
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: flin
 C:Superfamily: flagellar motor switch protein

Query Match 0.8%; Score 8; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSTYLEEL 20
 Db 27 LSTYLEEL 34

RESULT 5

H64592
 flagellar switch protein - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-Aug-1999
 C:Accession: H64592
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKen
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: H64592
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-123 <TOM>
 A:Cross-references: GB:AE000571; GB:AE000511; NID:G2313686; PIDN:AAD07636.1; PID:G2313686
 C:Superfamily: flagellar motor switch protein

Query Match 0.8%; Score 8; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSTYLEEL 20
 Db 27 LSTYLEEL 34

RESULT 6

B95264
 Probable lysoR-family protein [imported] - Sinorhizobium meliloti (strain 1021) magapias
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 27-Nov-2001
 C:Accession: B95264
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: B95264
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-313 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK64676.1; PID:G14523075; GSPDB:GN00165
 A:Experimental source: strain 1021, megapiasid pSYMA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMA0039
 A:Genome: plasmid
 C:Superfamily: conserved hypothetical protein H11364

Query Match 0.8%; Score 8; DB 2; Length 313;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 ELDLSFND 926
 Db 136 ELDLSFND 143

RESULT 7

S68209
 sds22 protein homolog - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C:Accession: S68209
 R:Renouf, S.; Beullens, M.; Wera, S.; Van Eynde, A.; Sikela, J.; Stalmans, W.; Bollen,
 FEBS Lett. 375, 75-78, 1995
 A:Title: Molecular cloning of a human polypeptide related to yeast sds22, a regulator
 A:Reference number: S68209; MUID:96087087; PMID:7498485
 A:Accession: S68209
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <REN>

A;Cross-references: EMBL:Z50749; NID:gl085027; PIDN:CAA90626.1; PID:gl085028
 C;Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 0.8%; Score 8; DB 2; Length 360;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 858 QSLRELDL 865
 Db 120 QSLRELDL 127
 |||||

RESULT 8
 S20597
 Ribonuclease inhibitor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
 C;Accession: S20597
 R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
 Biochim. Biophys. Acta 1129, 335-338, 1992
 A;Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
 A;Reference number: S20597; MUID:92162755; PMID:1536887
 A;Accession: S20597
 A;Molecule type: mRNA
 A;Residues: 1-456 <RAW>
 A;Cross-references: EMBL:X62528; NID:g57670; PIDN:CAA44388.1; PID:g57671
 C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 0.8%; Score 8; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 CRLRTLWL 837
 Db 251 CRLRTLWL 258
 |||||

RESULT 9
 A31857
 Ribonuclease inhibitor, hepatic - pig
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
 C;Accession: A31857; A35830
 R;Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
 Biochemistry 27, 8537-8544, 1988
 A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals th
 A;Reference number: A31857; MUID:89118268; PMID:3219361
 A;Accession: A31857
 A;Molecule type: protein
 A;Residues: 1-456 <HOF>
 R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
 Biochemistry 29, 8827-8834, 1990
 A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea
 A;Reference number: A35830; MUID:91104783; PMID:2271559
 A;Accession: A35830
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 82-456 <VIC>
 A;Cross-references: GB:M58700; NID:gl64638; PIDN:AAA63448.1; PID:gl64639; GB:J02925
 C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 C;Keywords: liver
 F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 0.8%; Score 8; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SURELDLS 866
 Db 394 SURELDLS 401
 |||||

RESULT 10

A31858
 Ribonuclease-angiogenin inhibitor - human
 N;Alternate names: ribonuclease inhibitor, placental
 C;Species: Homo sapiens (man)
 C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000
 C;Accession: A31858; S02012; S23933; S48636; T47188
 R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
 Biochemistry 27, 8545-8553, 1988
 A;Title: Primary structure of human placental ribonuclease inhibitor.
 A;Reference number: A31858; MUID:89118269; PMID:3219362
 A;Accession: A31858
 A;Molecule type: mRNA
 A;Residues: 1-461 <LEE>
 A;Cross-references: GB:M22414; NID:gl86260; PIDN:AA59130.1; PID:g307040
 R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
 EMBO J. 7, 4151-4156, 1988
 A;Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) disclose
 A;Reference number: S02012; MUID:89210799; PMID:3243277
 A;Accession: S02012
 A;Molecule type: mRNA
 A;Residues: 1-422, 'SE', 425-461 <SCH>
 A;Cross-references: EMBL:X13973; NID:g35843; PIDN:CAA32151.1; PID:g35844
 A;Note: part of this sequence, including the carboxyl end of the mature protein, was cor
 R;Crevel-thieffry, I.; Cottrell, S.; Schuller, E.
 Biochim. Biophys. Acta 1122, 107-112, 1992
 A;Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibit
 A;Reference number: S23933; MUID:92338217; PMID:1633192
 A;Accession: S23933
 A;Molecule type: protein
 A;Residues: 174-195 <CRE>
 R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.
 Arch. Biochem. Biophys. 312, 421-428, 1994
 A;Title: Purification and characterization of human brain ribonuclease inhibitor.
 A;Reference number: S48636; MUID:94311593; PMID:8037455
 A;Accession: S48636
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-14 <NAD>
 R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, March 2000
 A;Reference number: Z24374
 A;Accession: T47188
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-461 <AAA>
 A;Cross-references: EMBL:AL161967
 A;Experimental source: adult testis; clone DKFZp434K249
 C;Genetics:
 A;Gene: GDB:RNH
 A;Cross-references: GDB:125274; OMIM:173320
 A;Map position: lip15.5-lip15.5
 A;Note: DKFZp434K249.1
 C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 C;Keywords: blocked amino end; duplication
 F;34-433/Region: leucine-rich 57-residue repeats
 F;57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F;174-195/Region: inhibitory
 F;228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F;285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F;399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Query Match 0.8%; Score 8; DB 2; Length 461;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SURELDLS 866
 Db 399 SURELDLS 406
 |||||

RESULT 11

ZABPT4
helicase (EC 3.6.1.-) primase chain - phase T4
N/Alternate names: gp 41
C/Species: phage T4
C/Species: host Escherichia coli
C/Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 19-Jan-2001
C/Accession: A04308
R/Nakanishi, M.; Alberts, B.
unpublished results 1985, cited by GenBank
A/Reference number: A94456
A/Accession: A04308
A/Molecule type: DNA
A/Residues: 1-475 <NR>
A/Cross-references: GB:K03113; NID:g215964; PIDN:AAA32553.1; PID:g215966
R/Young, M.C.; Schultz, D.E.; Ring, D.; von Hippel, P.H.
J. Mol. Biol. 235, 1447-1458, 1994
A/Title: Kinetic parameters of the translocation of bacteriophage T4 gene 41 protein hel
A/Reference number: A58645; MUID:94149695; PMID:8107085
A/Contents: annotation; enzyme activity
C/Comment: This protein has GTPase, dGTPase and dATPase activities, and it is a helicase
C/Genetics:
A/Gene: 41
A/Map position: 20.619-22.044
A/Start codon: GUG
C/Superfamily: phage T4 gene 41 protein
C/Keywords: DNA binding; DNA repair; DNA replication; hydrolase; nucleotide binding; P-1
F:197-204/Region: nucleotide-binding motif A (P-loop)

Query Match 0.8%; Score 8; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 433 GLCSLAAD 440
|||||
Db 206 GLCSLAAD 213

RESULT 12
F86263
hypothetical protein Fl3K23.23 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C/Accession: F86263
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: F86263
A/Status: preliminary
A/Molecule type: DNA
A/Cross-references: GB:AE005172; NID:g8698744; PIDN:AAF78502.1; GSPDB:GN00141
C/Genetics:
A/Map position: 1

Query Match 0.8%; Score 8; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 NIRELDS 923
|||||
Db 377 NIRELDS 384

RESULT 13
T38422

probable chromatin assembly factor subunit - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
C/Accession: T38422
R/Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A/Reference number: Z21748
A/Accession: T38422
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-512 <OLI>
A/Cross-references: EMBL:Z99126; PIDN:CA16189.1; GSPDB:GN00066; SPDB:SPAC26H5.03
A/Experimental source: strain 972h-; cosmid c26H5
C/Genetics:
A/Gene: SPDB:SPAC26H5.03
A/Map position: 1
A/Introns: 44/3; 79/1; 160/1
C/Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 0.8%; Score 8; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 CLYEIQEE 596
|||||
Db 195 CLYEIQEE 202

RESULT 14
A11941
hypothetical protein all1084 [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: A11941
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: A11941
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-611 <KUR>
A/Cross-references: GB:BA000019; PIDN:BAB73041.1; PID:g17130430; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: all1084

Query Match 0.8%; Score 8; DB 2; Length 611;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 LDLSFNDL 927
|||||
Db 593 LDLSFNDL 600

RESULT 15
156258
RP105 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
C/Accession: 156258
R/Miyake, K.; Yamashita, Y.; Ogata, M.; Sudo, T.; Kimoto, M.
J. Immunol. 154, 3333-3340, 1995
A/Title: RP105, a novel B cell surface molecule implicated in B cell activation, is a m
A/Reference number: 156258; MUID:95204928; PMID:7897216
A/Accession: 156258
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-661 <RES>
A/Cross-references: GB:D37797; NID:g761711; PIDN:BAA07043.1; PID:g761712

C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match 0.8%; Score 8; DB 2; Length 661;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 NLRFLDLS 923
 |||||
 Db 371 NLRFLDLS 378

Search completed: July 30, 2004, 14:04:47
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 14:00:08 ; Search time 18 Seconds
(without alignments)
2994.034 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 1035

Sequence: 1 MLRTAGRDGLCRSLTYLEEL.....MTHSRLLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 141681 seqs, 52070155 residues
Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	818	79.0	1061	1 NAL12 HUMAN	P59046 homo sapien
2	13	1.3	1033	1 C1S1 MOUSE	Q8r4b8 mus musculus
3	13	1.3	1034	1 C1S1 HUMAN	Q96p20 homo sapien
4	12	1.2	1200	1 NAL5 HUMAN	P59047 homo sapien
5	9	0.9	980	1 NAL7 HUMAN	Q8wx94 homo sapien
6	9	0.9	1062	1 NAL2 HUMAN	Q9nx02 homo sapien
7	9	0.9	1093	1 NAL4 HUMAN	Q86w24 homo sapien
8	9	0.9	1473	1 NAL1 HUMAN	Q9c000 homo sapien
9	8	0.8	122	1 UCNI RAT	P55090 rattus norv
10	8	0.8	456	1 RINI PIG	P10775 sus scrofa
11	8	0.8	456	1 RINI RAT	P29315 rattus norv
12	8	0.8	460	1 RINI HUMAN	P13489 homo sapien
13	8	0.8	460	1 RINI PANTR	Q8hsp9 pan troglod
14	8	0.8	475	1 VG41 BPT4	P04530 bacterioph
15	8	0.8	661	1 C180 MOUSE	Q62192 mus musculus
16	8	0.8	707	1 IF2 THETN	Q8ra37 thermonaer
17	8	0.8	843	1 NAL6 MOUSE	Q91ws2 mus musculus
18	8	0.8	980	1 CLV1 ARATH	Q9syg8 arabidopsis
19	8	0.8	4036	1 RLPL DUGBV	Q66431 dugbe virus
20	7	0.7	54	1 RBL MAGLI	P31193 magnolia li
21	7	0.7	99	1 IHFA RHOC	P30787 rhodobacter
22	7	0.7	102	1 VE4 HPV1	P17384 human papil
23	7	0.7	116	1 YAP7 SHPO	Q09861 schizosacch
24	7	0.7	179	1 RUVC DEIRA	Q9rx75 delinococcus
25	7	0.7	193	1 PGCI PIG	Q95250 sus scrofa
26	7	0.7	194	1 PGCI MOUSE	O55022 mus musculus
27	7	0.7	194	1 PGCI RAT	P70580 rattus norv
28	7	0.7	195	1 ASC HUMAN	Q9ulz3 homo sapien
29	7	0.7	196	1 THIU ECOLI	Q46948 escherichia
30	7	0.7	196	1 THIU SALTY	P55880 salmonella
31	7	0.7	199	1 COX2 RHEAM	O03893 rhea ameri
32	7	0.7	224	1 KCY THETN	Q8ra78 thermonaer
33	7	0.7	238	1 ZNUC_BUCAI	P57403 buchnera ap

RESULT 1

ID	NAL12_HUMAN	STANDARD;	PRT;	1061 AA.
AC	P59046;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	NACHT-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like protein 7) (Monarch-1).			
DE	like protein 7) (Monarch-1).			
GN	NALP12 OR PYPAP7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=22162427; PubMed=12019269;			
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,			
RA	Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;			
RT	"PYPAP7, a novel PYRIN-containing Apaf1-like protein that regulates			
RT	activation of NF-kappa B and caspase-1-dependent cytokine			
RT	processing.";			
RL	J. Biol. Chem. 277:29874-29880(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22451042; PubMed=12563287;			
RA	Tschopp J., Martinon F., Burns K.;			
RT	"NALPs: a novel protein family involved in inflammation.";			
RL	Nat. Rev. Mol. Cell Biol. 4:95-104(2003).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).			
RC	TISSUE=Lymphoma;			
RA	Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Leukocyte;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			

ALIGNMENTS

P11402 spinacia ol
Q9by71 homo sapien
P59034 mus musculus
P59035 rattus norv
P25637 saccharomyc
Q9veb5 drosophila
Q02760 rhodobacter
Q55752 synechocyst
P33574 petromyzon
P39261 bacterioph
Q9ttt8 trichosurus
Q9m5j8 arabidopsis

34 7 0.7 257 1 ATPD SPIOL
35 7 0.7 257 1 LRR3 HUMAN
36 7 0.7 257 1 LRR3 MOUSE
37 7 0.7 257 1 LRR3 RAT
38 7 0.7 258 1 YCU9 YEAST
39 7 0.7 276 1 SRRI_DROME
40 7 0.7 285 1 CY1_RHOSH
41 7 0.7 321 1 RLPA SYNY3
42 7 0.7 325 1 PLMN PETMA
43 7 0.7 325 1 Y05F BPT4
44 7 0.7 328 1 GRHR TRIVU
45 7 0.7 330 1 PGI2_ARATH


```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/cvj;
RA Kikuchi-Yoshita R., Koga K., Taketomi Y., Sugiki T., Saico T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT "Identification of inducible genes during in vitro maturation of mouse
RT bone marrow-derived mast cells to connective tissue-type mast cells.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF486632; AAL90874.1; -.
DR MGD; MGI:2653833; C1asl.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_RL; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
DR Apoptosis; Repeat; Leucine-rich repeat.
FT DOMAIN 1 91
FT DAPIN.
FT REPEAT 216 532
FT REPEAT 737 760
FT REPEAT 794 817
FT REPEAT 851 874
FT REPEAT 880 903
FT REPEAT 908 931
FT REPEAT 937 964
FT REPEAT 965 988
FT REPEAT 1033 1033 AA; 118274 MW; 592469066B12117 CRC64;
SQ SEQUENCE 1033 AA; 118274 MW; 592469066B12117 CRC64;

Query Match 1.3%; Score 13; DB 1; Length 1033;
Best Local Similarity 100.0%; Pred.No. 0.00018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 HPRHVEILGFSEA 341
Db 360 HPRHVEILGFSEA 372
|||||
|||||

RESULT 3
CIS1_HUMAN STANDARD; PRT; 1034 AA.
AC Q96P20; O75434; Q8TCW0; Q8TEU9; Q8WXH9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cold autoimmune syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
DE PYD-containing protein 3) (PRIN-containing APAF1-like protein 1)
DE (Angiotensin/vasopressin receptor AII/AVP-like).
GN CIAS1 OR NALP3 OR PYPAF1.
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2); VARIANTS FCAS MET-198; VAL-439
RP AND GLY-627; AND VARIANT MWS VAL-352.
RX MEDLINE=21547523; PubMed=11687797;
RA Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A.,
RA Kolodner R.D.;
RT "Mutation of a new gene encoding a putative pyrin-like protein causes
RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome.";
RL Nat. Genet. 29:301-305(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21909508; PubMed=11786556;
RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
RA Mak S., Lora J.M., Briskin M., Jurman M., Cao J., DiStefano P.S.,
RA Bertin J.;
RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
RT and activates NF-kB.";
RL J. Biol. Chem. 277:11570-11575(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3); VARIANT MWS MET-198, AND
RP VARIANTS FCAS TRP-260 AND PRO-305.
RX MEDLINE=22241234; PubMed=12355493;
RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
RA Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
RA Cotter F.E., Thome M., Hitman G.A., Tschopp J., McDermott M.F.;
RT "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad
RT phenotype including recurrent fever, cold sensitivity, sensorineural
RT deafness, and AA amyloidosis.";
RL Arthritis Rheum. 46:2445-2452(2002).
RN [4]
RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
RC TISSUE=Blood;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
RN [5]
RP VARIANTS FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND
RP ARG-569, AND VARIANT FCAS/MWS TRP-260.
RX MEDLINE=21987640; PubMed=11992256;
RA Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,
RA Granel B., Frances C., Garcier F., Edery P., Boulanguet S.,
RA Domergues J.-P., Delpech M., Grateau G.;
RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
RT and familial cold urticaria: a novel mutation underlies both
RT syndromes.";
RL Am. J. Hum. Genet. 70:1498-1506(2002).
RN [6]
RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
RP THR-662, AND TISSUE SPECIFICITY.
RX MEDLINE=22062556; PubMed=12032915;
RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
RT "Chronic infantile neurological cutaneous and articular syndrome is
RT caused by mutations in CIAS1, a gene highly expressed in
RT polymorphonuclear cells and chondrocytes.";
RL Am. J. Hum. Genet. 71:198-203(2002).
CC -!- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC IsoId=Q96P20-1; Sequence=Displayed;

```


RT for early embryonic development.";
 RL Hum. Reprod. 17:903-911(2002).
 CC -!- TISSUE SPECIFICITY: Oocyte specific.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -!- CAUTION: It is not obvious that this is the ortholog of mouse Mater.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AY054986; AAL15549.1; -;
 CC Genew; HGNC:21269; NALP5.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR007091; LRR RNinh.
 CC InterPro; IPR003590; LRR RNinh_sub.
 CC InterPro; IPR007111; NACHT_NTPase.
 CC InterPro; IPR004020; PAAD_DAPIN_dom.
 CC Pfam; PF00560; LRR: 3.
 CC Pfam; PF05729; NACHT; 1.
 CC Pfam; PF02758; PAAD_DAPIN; 1.
 CC SMART; SM00368; LRR_R1; 11.
 CC PROSITE; PS50824; DAPIN; 1.
 CC PROSITE; PS50837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 57 148 DAPIN.
 FT DOMAIN 280 602 NACHT.
 FT REPEAT 704 727 LRR 1.
 FT REPEAT 730 753 LRR 2.
 FT REPEAT 780 803 LRR 3.
 FT REPEAT 809 832 LRR 4.
 FT REPEAT 836 863 LRR 5.
 FT REPEAT 865 892 LRR 6.
 FT REPEAT 893 916 LRR 7.
 FT REPEAT 950 973 LRR 8.
 FT REPEAT 979 1002 LRR 9.
 FT REPEAT 1007 1034 LRR 10.
 FT REPEAT 1036 1059 LRR 11.
 FT REPEAT 1064 1092 LRR 12.
 FT REPEAT 1121 1142 LRR 13.
 FT NP_BIND 286 293 ATP (POTENTIAL).
 SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A771B28FA CRC64;
 Query Match 1.2%; Score 12; DB 1; Length 1200;
 Best Local Similarity 100.0%; Pred. No. 0.0022;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 259 PERLLFIIDGFD 270
 Db 354 PERLLFIIDGFD 365
 |||||
 RESULT 5
 ID NAL7 HUMAN STANDARD; PRT; 980 AA.
 AC Q8WY94;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 7 (PYRIN-containing APAF1-like
 DE protein 3).
 GN NALP7 OR PYPAF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing.";
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22451042; PubMed=12563287;
 RA Tschopp J., Martinon F., Burns K.;
 RT "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF464765; AAL69963.1; -;
 CC EMBL; AY154462; AAO18158.1; -;
 CC Genew; HGNC:22947; NALP7.
 CC InterPro; IPR007091; LRR RNinh.
 CC InterPro; IPR007111; NACHT_NTPase.
 CC InterPro; IPR004020; PAAD_DAPIN_dom.
 CC Pfam; PF05729; NACHT; 1.
 CC Pfam; PF02758; PAAD_DAPIN; 1.
 CC PROSITE; PS50824; DAPIN; 1.
 CC PROSITE; PS50837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 93 DAPIN.
 FT DOMAIN 172 491 NACHT.
 FT REPEAT 614 638 LRR 1.
 FT REPEAT 674 697 LRR 2.
 FT REPEAT 760 784 LRR 3.
 FT REPEAT 788 810 LRR 4.
 FT REPEAT 817 840 LRR 5.
 FT REPEAT 845 868 LRR 6.
 FT REPEAT 874 897 LRR 7.
 FT REPEAT 902 928 LRR 8.
 FT REPEAT 933 957 LRR 9.
 FT NP_BIND 178 185 ATP (POTENTIAL).
 SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;
 Query Match 0.9%; Score 9; DB 1; Length 980;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 483 YSFIHLSFQ 491
 Db 464 YSFIHLSFQ 472
 |||||
 RESULT 6
 ID NAL2 HUMAN STANDARD; PRT; 1062 AA.
 AC Q9NX02; Q9BVN5; Q9H6G6; Q9HAV9; Q9NWK3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 2 (PYRIN-containing APAF1-like
 DE protein 2) (Nucleotide-binding site protein 1).
 GN NALP2 OR PYPAF2 OR NBS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CC NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 CC SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., DiStefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 proteins.";
 RL Cell Death Differ. 7:1273-1274(2000).
 RN [2]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 implicated in apoptosis and inflammation.";
 RL Curr. Biol. 11:R118-R120(2001).
 RN [3]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 activation of NF-kappa B and caspase-1-dependent cytokine
 processing.";
 RL J. Biol. Chem. 277:29874-29880(2002).
 RN [4]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: May be implicated in apoptosis (By similarity).
 CC -1- COFACTOR: Binds ATP (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9NX02-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NX02-2; Sequence=VSP_005522;
 CC -1- SIMILARITY: Contains 1 DAPIN domain.
 CC -1- SIMILARITY: Contains 1 NACHT domain.
 CC -1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AF298547; AAG15253.1; ALT_INIT.
 CC EMBL; AF310106; AAG30289.1; -
 CC EMBL; AF464764; AAL69962.1; -
 CC EMBL; AK000517; BAA91223.1; -
 CC EMBL; AK025952; BAB15293.1; -
 CC EMBL; AK000784; BAA91377.1; ALT_INIT.
 CC EMBL; BC003592; AAH03592.1; -
 CC EMBL; BC001039; AAH01039.1; -
 CC HSSP; PI0775; 2BNH.
 CC Genew; HGNC:22948; NALP2.
 CC InterPro; IPR007091; LRR_RNinh.
 CC InterPro; IPR007111; NACHT_NTPase.
 CC InterPro; IPR004020; PAAD_DAPIN_dom.
 CC Pfam; PF05729; NACHT; 1.
 CC PROSITE; PS0824; DAPIN; 1.
 CC PROSITE; PS0837; NACHT; 1.
 CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 94 DAPIN.
 FT DOMAIN 207 526 NACHT.
 FT REPEAT 467 491 LRR 1.
 FT REPEAT 622 645 LRR 2.
 FT REPEAT 754 777 LRR 3.
 FT REPEAT 810 832 LRR 4.
 FT REPEAT 839 862 LRR 5.
 FT REPEAT 867 890 LRR 6.
 FT REPEAT 924 947 LRR 7.
 FT REPEAT 981 1005 LRR 8.
 FT REPEAT 1010 1033 LRR 9.
 FT NP_BIND 213 220 ATP (POTENTIAL).
 FT DOMAIN 518 523 POLY-GLU.
 FT VARSPLIC 133 154 Missing (in isoform 2).
 FT CONFLICT 1 1 M -> V (IN REF. 1).
 FT CONFLICT 35 35 L -> P (IN REF. 1).
 FT CONFLICT 304 304 I -> S (IN REF. 1).
 FT CONFLICT 364 364 R -> K (IN REF. 1).
 FT CONFLICT 980 980 MISSING (IN REF. 1).
 FT CONFLICT 1052 1052 A -> E (IN REF. 5).
 SQ SEQUENCE 1062 AA; 4DBB0F6E9C2BC8A7 CRC64;
 Query Match 0.9%; Score 9; DB 1; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 483 YSFHLSFQ 491
 Db 499 YSFHLSFQ 507
 RESULT 7
 NA14 HUMAN STANDARD; PRT; 1093 AA.
 ID NA14 HUMAN
 AC Q86W24;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 14.
 GN NALP14.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22451042; PubMed=12563287;
 RA Tschopp J., Martinon F., Burns K.;
 RT "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).

CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AV154469; AAO18165.1; -;
 DR Genew; HGNC:22339; NALP14.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR RNinh.
 DR InterPro; IPR003590; LRR RNinh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00019; LEURICRPT.
 DR SMART; SM00368; LRR RT; 12.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR ATP-binding; Leucine-rich repeat; Repeat.

KW DOMAIN 1 97 DAPIN.
 FT 177 499 NACHT.
 FT REPEAT 525 550 LRR 1.
 FT REPEAT 672 695 LRR 2.
 FT REPEAT 733 756 LRR 3.
 FT REPEAT 757 780 LRR 4.
 FT REPEAT 785 807 LRR 5.
 FT REPEAT 842 869 LRR 6.
 FT REPEAT 871 894 LRR 7.
 FT REPEAT 899 926 LRR 8.
 FT REPEAT 956 979 LRR 9.
 FT REPEAT 985 1008 LRR 10.
 FT REPEAT 1013 1037 LRR 11.
 SQ SEQUENCE 1093 AA; 124732 MW; 124EAFCE22A11D6F CRC64;

Query Match 0.9%; Score 9; DB 1; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 2.5; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

QY 711 LRHPNCKLQ 719
 |||||
 Db 697 LRHPNCKLQ 705

RESULT 8
 ID NAL1 HUMAN STANDARD; PRT; 1473 AA.
 AC Q9C000; Q9BZ28; Q9HAV8; Q9UFT4; Q9Y2B0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-
 DE forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 DE caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA "Hertin J., DiStefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 RT proteins.";
 RL Cell Death Differ. 7:1273-1274(2000).

RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 RL implicated in apoptosis and inflammation.";
 RL Curr. Biol. 11:R118-R120(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brythroleukemia;
 RX MEDLINE=21153743; PubMed=11076957;
 RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
 RA Shi M.M., Vincenz C., Ward P.A.;
 RT "Molecular cloning and characterization of DEFCAP-L and -S, two
 RT isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 RT proteins.";
 RL J. Biol. Chem. 276:9230-9238(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.
 RC TISSUE=T-cell;
 RX MEDLINE=21153744; PubMed=11113115;
 RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
 RA Godzik A., Reed J.C.;
 RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
 RT c-dependent caspase activation and apoptosis.";
 RL J. Biol. Chem. 276:9239-9245(2001).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RL for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [6]
 RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
 RC TISSUE=Uterus;
 RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Able to form cytoplasmic structures termed death
 CC effector filaments. Enhances APAF1 and cytochrome c-dependent
 CC activation of pro-caspase-9 and consecutive apoptosis. Seems to
 CC bind ATP.
 CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
 CC and with APAF1 in a cytochrome c-inducible way leading to the
 CC formation of an apoptosome. This interaction may be ATP-dependent.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=1; Synonyms=NAC beta, DEFCAP-L;
 CC IsoId=Q9C000-1; Sequence=Displayed;
 CC Name=2; Synonyms=NAC alpha, DEFCAP-S;
 CC IsoId=Q9C000-2; Sequence=VSP_004327;
 CC Name=3; Synonyms=NAC gamma;
 CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
 CC Name=4; Synonyms=NAC delta;
 CC IsoId=Q9C000-4; Sequence=VSP_004326;
 CC -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
 CC expressed in peripheral blood leukocytes, chronic myelogenous
 CC leukemia cell line K-562, followed by thymus, spleen and heart.
 CC Also detected in lung, placenta, small intestine, colon, kidney,
 CC liver and muscle.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 1 CARD domain.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

-----
CC EMBL; AF298548; AAG15254.1; -
CC EMBL; AF310105; AAG30288.1; -
CC EMBL; AF229059; AAK00748.1; -
CC EMBL; AF229060; AAK00749.1; -
CC EMBL; AF229061; AAK00750.1; -
CC EMBL; AF229062; AAK00751.1; -
CC EMBL; AB023143; BAA00770.1; -
CC EMBL; AL117470; CAB55945.1; -
CC PIR; T17255; T17255.
CC HSP; P13489; IAA4Y.
CC MIM; 606636; -.
CC GO; GO:0005622; C:intracellular; IC.
CC GO; GO:0016506; F:apoptosis activator activity; NAS.
CC GO; GO:0008566; F:caspase activator activity; NAS.
CC GO; GO:0019899; F:enzyme binding; IPI.
CC GO; GO:0006919; F:caspase activation; NAS.
CC GO; GO:0006917; F:induction of apoptosis; NAS.
CC InterPro; IPR001315; CARD.
CC InterPro; IPR000767; Disease_resist.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_Rninh.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF00560; LRR; 2.
CC Pfam; PF05729; NACHT; 1.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC PRINTS; PR00364; DISEASERIST.
CC PROSITE; PS0209; CARD; 1.
CC PROSITE; PS00824; DAPIN; 1.
CC PROSITE; PS00837; NACHT; 1.
CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW Alternative splicing.
FT DOMAIN 1 92
FT DAPIN.
FT DOMAIN 328 637
FT NACHT.
FT REPEAT 704 725
FT LRR 1.
FT REPEAT 807 830
FT LRR 2.
FT REPEAT 864 887
FT LRR 3.
FT REPEAT 921 944
FT LRR 4.
FT REPEAT 950 973
FT LRR 5.
FT REPEAT 1199 1215
FT LRR 6.
FT REPEAT 1216 1236
FT LRR 7.
FT DOMAIN 1374 1463
FT CARD.
FT NP BIND 334 341
FT VARSP LIC 958 987
FT Missing (in isoform 3 and isoform 4).
FT /FTid=VSP_004326.
FT Missing (in isoform 2 and isoform 3).
FT /FTid=VSP_004327.
FT K->L: NO EFFECT.
FT K->S: NO EFFECT.
FT L->H (IN REF. 1).
FT T->S (IN REF. 1).
FT T->S (IN REF. 1).
FT T->M (IN REF. 1).
FT T->I (IN REF. 1).
FT M->V (IN REF. 1).
FT M->V (IN REF. 1 AND 6).
FT V->L (IN REF. 1).
FT R->C (IN REF. 1).
FT SEQUENCE 1473 AA; 165865 MW; 438F0DCE45C2562D CRC64;

```

Query Match 0.9%; Score 9; DB 1; Length 1473;
 Best Local Similarity 100.0%; Pred.No. 3.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 QGAGIGKGS 197
 Db 333 QGAGIGKGS 341

RESULT 9

```

UC1 RAT STANDARD; PRT; 122 AA.
AC P55030;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Urocortin precursor (Corticotensin).
GN UCN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96069764; PubMed=7477349;
RA Vaughan J.M., Donaldson C.J., Bittencourt J., Perrin M.H., Lewis K.A.,
RA Sutton S.W., Chan R., Turnbull A., Lovejoy D., Rivier C., Rivier J.E.,
RA Sawchenko P., Vale W.W.;
RA "Urocortin, a mammalian neuropeptide related to fish urotensin I and
RT to corticotropin-releasing factor.";
RL Nature 378:287-292 (1995).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=Lewis;
RA Park J.H., Lee Y.J., Kim K.L.;
RA "Detection of rat urocortin in lymphoid tissues: implications for the
RT functional assessment of urocortin as a novel neuro-immunomodulatory
RT peptide.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts in vitro to stimulate the secretion of
CC adrenocorticotrophic hormone (ACTH). Binds with high affinity to
CC CRF Receptor types 1, 2-alpha, and 2-beta.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing
CC factor/urotensin I family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

EMBL; U33935; AAA87566.1; -
 EMBL; AF093623; AAF63153.1; -
 PIR; S60262; S60262.
 InterPro; IPR000187; corticoliberin.
 InterPro; IPR003620; Urocortin_CRF.
 Pfam; PF00473; CRF; 1.
 PRINTS; PR01612; CRFFAMILY.
 SMART; SMO0039; CRF; 1.
 PROSITE; PS00511; CRF; 1.
 Hormone; Amidation; Cleavage on pair of basic residues; Signal.
 SIGNAL 1 25
 PROPEP 26 80
 FT PEPTIDE 81 120
 FT MOD_RES 120 120
 FT AMIDATION (G-121 PROVIDE AMIDE GROUP) (BY
 FT SIMILARITY).
 SQ SEQUENCE 122 AA; 13711 MW; 9F0AF834CBFCF74 CRC64;

Query Match 0.8%; Score 8; DB 1; Length 122;
 Best Local Similarity 100.0%; Pred.No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 LLVQLRPE 667
 Db 15 LLVQLRPE 22

RESULT 10


```

RINI_PIG
ID RINI_PIG STANDARD; PRT; 456 AA.
AC 010775;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease inhibitor.
GN RNH OR RI.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE
RC TISSUE=Liver;
RX MEDLINE=89118268; PubMed=3219361;
RA Hofsteenge J., Kieffer B., Matthies R., Hemmings B.A., Stone S.R.;
RT "Amino acid sequence of the ribonuclease inhibitor from porcine liver
RT reveals the presence of leucine-rich repeats.";
RL Biochemistry 27:8537-8544(1988).
RN [2]
RP SEQUENCE OF 82-456 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=91104783; PubMed=2271559;
RA Vicentini A.M., Kieffer B., Matthies R., Meyhack B., Hemmings B.A.,
RA Stone S.R., Hofsteenge J.;
RT "Protein chemical and kinetic characterization of recombinant porcine
RT ribonuclease inhibitor expressed in Saccharomyces cerevisiae.";
RL Biochemistry 29:8827-8834(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=94088748; PubMed=8264799;
RA Kobe B., Deisenhofer J.;
RT "Crystal structure of porcine ribonuclease inhibitor, a protein with
RT leucine-rich repeats.";
RL Nature 366:751-756(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=95183144; PubMed=7877692;
RA Kobe B., Deisenhofer J.;
RT "A structural basis of the interactions between leucine-rich repeats
RT and protein ligands.";
RL Nature 374:183-186(1995).
CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
CC function in the modulation of cellular activities.
CC -!- SUBUNIT: forms a tight one-to-one complex with the RNase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M58700; AAA63448.1; -.
CC F01; A31857; A31857.
CC PDB; 2NH; 12-MAR-97.
CC PDB; 1DFJ; 11-JAN-97.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_Rninh.
CC InterPro; IPR003590; LRR_Rninh_sub.
CC Pfam; PF00560; LRR; 5.
CC SMART; SM00368; LRR_RI; 2.
KW Acetylation; Repeat; Leucine-rich repeat; 3D-structure.
FT REPEAT 15 43 LRR A1.
FT REPEAT 44 71 LRR B1.
FT REPEAT 72 100 LRR A2.
FT REPEAT 101 128 LRR B2.
FT REPEAT 129 157 LRR A3.
FT REPEAT 158 185 LRR B3.

```

```

FT REPEAT 214 214 LRR A4.
FT REPEAT 215 242 LRR B4.
FT REPEAT 243 271 LRR A5.
FT REPEAT 272 299 LRR B5.
FT REPEAT 300 328 LRR A6.
FT REPEAT 329 356 LRR B6.
FT REPEAT 357 385 LRR A7.
FT REPEAT 386 413 LRR B7.
FT REPEAT 414 442 LRR A8.
FT MOD RES 1 1 ACETYLATION.
FT STRAND 2 2
FT STRAND 5 6
FT HELIX 12 22
FT TURN 23 24
FT STRAND 27 31
FT TURN 32 32
FT HELIX 37 47
FT TURN 48 49
FT TURN 51 52
FT STRAND 55 57
FT HELIX 64 75
FT TURN 78 79
FT STRAND 84 86
FT TURN 88 89
FT HELIX 94 98
FT HELIX 100 106
FT TURN 108 109
FT STRAND 112 114
FT HELIX 121 133
FT TURN 135 136
FT STRAND 141 143
FT TURN 145 146
FT STRAND 150 150
FT STRAND 169 171
FT TURN 173 174
FT STRAND 176 176
FT HELIX 178 180
FT TURN 191 191
FT STRAND 198 200
FT TURN 202 203
FT HELIX 208 220
FT TURN 222 223
FT STRAND 226 228
FT HELIX 235 246
FT TURN 247 247
FT TURN 249 250
FT STRAND 255 257
FT TURN 259 260
FT HELIX 265 277
FT STRAND 283 285
FT TURN 287 288
FT HELIX 292 303
FT TURN 304 304
FT TURN 306 307
FT STRAND 312 314
FT TURN 316 317
FT STRAND 321 321
FT HELIX 322 324
FT STRAND 325 334
FT STRAND 340 342
FT STRAND 347 347
FT HELIX 349 359
FT TURN 360 361
FT STRAND 369 371
FT TURN 373 374
FT HELIX 379 391
FT STRAND 397 399
FT TURN 401 402
FT HELIX 407 417
FT TURN 418 418
FT STRAND 426 428
FT TURN 430 431

```

```
FT HELIX 436 448
FT STRAND 453 456
SQ SEQUENCE 456 AA; 49023 MW; 01DA0A529CDC763E CRC64;

Query Match 0.8%; Score 8; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SRELDLS 866
Db 394 SRELDLS 401

RESULT 11
RINI_RAT
ID RINI_RAT STANDARD; PRT; 456 AA.
AC P29315;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease inhibitor.
DE RNH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=92162755; PubMed=1536887;
RA Kawamoto M., Motojima K., Sasaki M., Hattori H., Goto S.;
RT "cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue
distribution of the mRNA.";
RL Biochim. Biophys. Acta 1129:335-338(1992).
CC -1- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
function in the modulation of cellular activities.
CC -1- SUBUNIT: Forms a tight one-to-one complex with the RNase.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: Brain, heart, lung, liver, spleen, testes and
kidney; highest in the lung and lowest in the heart.
CC -1- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X62528; CAA4388.1; --
CC F1R; S20597; S20597.
CC HSSP; P10775; 2BNH.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_Rninh.
CC Pfam; PF00560; LRR_4.
CC SMART; SM00368; LRR_R1; 1.
KW Repeat; Leucine-rich repeat.
FT REPEAT 15 43 LRR A1.
FT REPEAT 44 71 LRR B1.
FT REPEAT 72 100 LRR A2.
FT REPEAT 101 128 LRR B2.
FT REPEAT 129 157 LRR A3.
FT REPEAT 158 185 LRR B3.
FT REPEAT 186 214 LRR A4.
FT REPEAT 215 242 LRR A5.
FT REPEAT 243 271 LRR B5.
FT REPEAT 272 299 LRR A6.
FT REPEAT 300 328 LRR B6.
FT REPEAT 329 356 LRR A7.
FT REPEAT 357 385 LRR A8.
FT REPEAT 386 413 LRR B7.

FT REPEAT 414 442 LRR A8.
SQ SEQUENCE 456 AA; 49905 MW; 8518E5B1F09E5998 CRC64;

Query Match 0.8%; Score 8; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 CRLRTLWL 837
Db 251 CRLRTLWL 258

RESULT 12
RINI_HUMAN
ID RINI_HUMAN STANDARD; PRT; 460 AA.
AC P13489;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental ribonuclease inhibitor (Ribonuclease/angiogenin inhibitor)
DE (RAI) (RNase inhibitor) (RI).
DE RNH OR PRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118269; PubMed=3219362;
RA Lee F.S., Fox E.A., Zhou H.-M., Strydom D.J., Vallee B.L.;
RT "Primary structure of human placental ribonuclease inhibitor.";
RL Biochemistry 27:8545-8553(1988).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 443-461.
RX MEDLINE=89210799; PubMed=3243277;
RA Schneider R., Schneider-Scherzer E., Thurnher M., Auer B.,
RA Schweiger M.;
RT "The primary structure of human ribonuclease/angiogenin inhibitor
(RAI) discloses a novel highly diversified protein superfamily with a
common repetitive module.";
RL EMBO J. 7:4151-4156(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Kidney, Lymph, and Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF COMPLEX WITH ANGIOGENIN.
RX MEDLINE=97459904; PubMed=9311977;
```

RA Papageorgiou A.C., Shapiro R., Acharya K.R.;
 RT "Molecular recognition of human angiogenin by placental ribonuclease
 RL inhibitor: -- an X-ray crystallographic study at 2.0-A resolution.";
 CC EMBO J. 16:5162-5177(1997).
 CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
 CC function in the modulation of cellular activities.
 CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; M22414; AAA59130.1; -;
 CC EMBL; X13973; AAA32151.1; -;
 CC EMBL; M36717; AAA60249.1; -;
 CC EMBL; AL161967; CAB82310.1; -;
 CC EMBL; BC003075; AAH03075.1; -;
 CC EMBL; BC011500; AAH11500.1; -;
 CC EMBL; BC014629; AAH14629.1; -;
 CC EMBL; BC047730; AAH47730.1; -;
 CC PIR; A31858; A31858.
 CC PDB; 1A4Y; 14-OCT-98.
 CC Genew; HGNC:10074; RNH.
 CC MIM; 173320; -;
 CC GO; GO:0008428; F:ribonuclease inhibitor activity; TAS.
 CC GO; GO:006401; P:RNA catabolism; TAS.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR007091; LRR_RNinh.
 CC InterPro; IPR003590; LRR_RNinh_sub.
 CC Pfam; PF00560; LRR; 3.
 CC SMART; SM00368; LRR_R1; 1.
 KW Repeat; Leucine-rich repeat; 3D-structure; Polymorphism.
 FT INIT_MET 0
 FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
 FT REPEAT 19 47 LRR A1.
 FT REPEAT 48 75 LRR B1.
 FT REPEAT 76 104 LRR A2.
 FT REPEAT 105 132 LRR B2.
 FT REPEAT 133 161 LRR A3.
 FT REPEAT 162 189 LRR B3.
 FT REPEAT 190 218 LRR A4.
 FT REPEAT 219 246 LRR B4.
 FT REPEAT 247 275 LRR A5.
 FT REPEAT 276 303 LRR B5.
 FT REPEAT 304 332 LRR A6.
 FT REPEAT 333 360 LRR B6.
 FT REPEAT 361 389 LRR A7.
 FT REPEAT 390 417 LRR B7.
 FT REPEAT 418 446 LRR A8.
 FT VARIANT 169 169 P -> L (in dbSNP:17585).
 FT
 FT CONFLICT 422 423 /FTId-VAR_014726.
 FT STRAND 2 10 RQ -> SE (IN REF. 2).
 FT HELIX 16 22
 FT TURN 23 28
 FT STRAND 31 35
 FT TURN 41 43
 FT HELIX 44 52
 FT TURN 53 53
 FT STRAND 59 61
 FT TURN 63 64
 FT HELIX 68 76
 FT TURN 77 79
 FT TURN 82 83
 FT STRAND 88 90
 FT TURN 92 93
 FT HELIX 98 100

FT HELIX 101 110
 FT TURN 112 113
 FT STRAND 116 118
 FT HELIX 125 136
 FT TURN 137 137
 FT TURN 139 140
 FT STRAND 145 147
 FT TURN 149 150
 FT HELIX 155 157
 FT HELIX 158 167
 FT TURN 169 170
 FT STRAND 173 175
 FT HELIX 182 195
 FT STRAND 202 204
 FT TURN 206 207
 FT TURN 213 213
 FT HELIX 214 224
 FT TURN 226 227
 FT STRAND 230 232
 FT HELIX 239 250
 FT TURN 251 251
 FT TURN 253 254
 FT STRAND 259 261
 FT TURN 263 264
 FT HELIX 269 279
 FT TURN 280 281
 FT TURN 283 284
 FT STRAND 287 289
 FT TURN 291 292
 FT HELIX 296 307
 FT TURN 308 308
 FT TURN 310 311
 FT STRAND 316 318
 FT TURN 320 321
 FT STRAND 325 325
 FT HELIX 326 328
 FT HELIX 329 338
 FT STRAND 344 346
 FT STRAND 351 351
 FT HELIX 353 364
 FT TURN 365 365
 FT TURN 367 368
 FT STRAND 373 375
 FT TURN 377 378
 FT TURN 383 385
 FT HELIX 386 395
 FT STRAND 401 403
 FT TURN 405 406
 FT HELIX 411 422
 FT TURN 424 425
 FT STRAND 430 432
 FT TURN 434 435
 FT HELIX 440 452
 FT TURN 454 455
 FT STRAND 457 460
 SQ SEQUENCE 460 AA; 49842 MW; C3D6668E2F2BPF86F CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 460;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 859 SIRELDLS 866
 Db 398 SIRELDLS 405
 |||||
 RESULT 13
 ID RINI_PANTR STANDARD; PRT; 460 AA.
 AC Q8HZE9;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Placental ribonuclease inhibitor (Ribonuclease/angiogenin inhibitor)
 DE (RAI) (RNase inhibitor) (RI).
 GN RNH.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Zhang J., Webb D.M., Podlaha O.;
 RT "Accelerated protein evolution and origins of human-specific features:
 RT FOX2 as an example.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
 CC function in the modulation of cellular activities (By similarity).
 CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF539549; AAN10133.1; -.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR006553; LRR_CYS_sub.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR Pfam; PF00560; LRR; 3.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00367; LRR_CC; 5.
 DR SMART; SM00368; LRR_RI; 13.
 DR Repeat; Leucine-rich repeat.
 KW INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
 FT REPEAT 19 47 LRR A1.
 FT REPEAT 48 75 LRR B1.
 FT REPEAT 76 104 LRR A2.
 FT REPEAT 105 132 LRR B2.
 FT REPEAT 133 161 LRR A3.
 FT REPEAT 162 189 LRR B3.
 FT REPEAT 190 218 LRR A4.
 FT REPEAT 219 246 LRR B4.
 FT REPEAT 247 275 LRR A5.
 FT REPEAT 276 303 LRR B5.
 FT REPEAT 304 332 LRR A6.
 FT REPEAT 333 360 LRR B6.
 FT REPEAT 361 389 LRR A7.
 FT REPEAT 390 417 LRR B7.
 FT REPEAT 418 446 LRR A8.
 SQ SEQUENCE 460 AA; 49798 MW; C11F15EE97DBB4F CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 460;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 859 SLRELDLS 866
 DB 398 SLRELDLS 405
 RESULT 14
 VC41 BPT4
 ID VC41 BPT4 STANDARD; PRT; 475 AA.
 AC P04530;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Primase-helicase (Protein Gp41).
 GN 41.
 OS Bacteriophage T4.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=10665;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=89340565; PubMed=2668290;
 RA Nakanishi M., Alberts B.;
 RL Unpublished results, cited by:
 RL Hinton D.M.;
 RL J. Biol. Chem. 264:14440-14446(1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=22514363; PubMed=12626685;
 RA Miller E.S., Kutter E., Mosig G., Arisaka F., Kunisawa T., Ruger W.;
 RT "Bacteriophage T4 genome.";
 RL Microbiol. Mol. Biol. Rev. 67:96-156(2003).
 RN [3]
 RN CHARACTERIZATION.
 RP MEDLINE=94149695; PubMed=8107085;
 RP Young M.C., Schlutz D.E., Ring D., von Hippel P.H.;
 RA "Kinetic parameters of the translocation of bacteriophage T4 gene 41
 RT protein helicase on single-stranded DNA.";
 RL J. Mol. Biol. 235:1447-1458(1994).
 CC -!- FUNCTION: Essential replication protein, part of the primase-
 CC helicase required for lagging strand DNA synthesis. It acts
 CC processively. It forms a protein complex with the gene 59 protein
 CC that partly replaces the dda protein helicase function. Act as
 CC single-stranded ATP-dependent DNA helicase.
 CC -!- MISCELLANEOUS: Interacts with the gene 61 protein to form the T4
 CC primosome.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; K03113; AAA32553.1; -.
 DR EMBL; AF158101; AAD42466.1; -.
 DR PIR; A04308; Z4BPT4.
 DR InterPro; IPR007694; DnaB_C.
 DR Pfam; PF03796; DnaB_C_1.
 DR Helicase; DNA replication; ATP-binding; DNA-binding.
 KW NP BIND 197 204 ATP (POTENTIAL).
 FT SEQUENCE 475 AA; 53601 MW; A544B1F9CFC90B4C CRC64;
 SQ
 Query Match 0.8%; Score 8; DB 1; Length 475;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 433 GLCSLAAD 440
 DB 206 GLCSLAAD 213
 RESULT 15
 C180 MOUSE
 ID C180 MOUSE STANDARD; PRT; 661 AA.
 AC Q62192;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE CD180 antigen precursor (Lymphocyte antigen 78) (Radioprotective 105
 DE kDa protein).
 DE LY78 OR RP105.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 21-43.
STRAIN=BALB/c; TISSUE=B-cell lymphoma;
MEDLINE=95204928; PubMed=7897216;
RA Miyake K., Yamashita Y., Ogata M., Sudo T., Kimoto M.;
RT "RP105, a novel B cell surface molecule implicated in B cell
RL activation, is a member of the leucine-rich repeat protein family."
J. Immunol. 154:3333-3340(1995).
[2]
INTERACTION WITH MD-1.
RA MEDLINE=98349386; PubMed=9686597;
RA Miyake K., Shimazu R., Kondo J., Niki T., Akashi S., Ogata H.,
RA Yamashita Y., Miura Y., Kimoto M.;
RT "Mouse MD-1, a molecule that is physically associated with RP105 and
RL positively regulates its expression."
J. Immunol. 161:1348-1353(1998).
[3]
FUNCTION.
RX MEDLINE=20341616; PubMed=10880523;
RA Ogata H., Su I., Miyake K., Nagai Y., Akashi S., Mecklenbrauer I.,
RA Rajewsky K., Kimoto M., Tarakhovskiy A.;
RT "The Toll-like receptor protein RP105 regulates lipopolysaccharide
RL signaling in B cells."
J. Exp. Med. 192:23-29(2000).
CC -!- FUNCTION: May cooperate with MD-1 and TLR4 to mediate the innate
CC immune response to bacterial lipopolysaccharide (LPS) in B cells.
CC Leads to NF-kappa-B activation. Also involved in the life/death
CC decision of B cells.
CC -!- SUBUNIT: Binds to MD-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein; plasma membrane.
CC -!- TISSUE SPECIFICITY: B lymphocytes and spleen. Not detected in
CC thymus, kidney, muscle, heart, brain or liver.
CC -!- SIMILARITY: Belongs to the Toll-like receptor family.
CC -!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D37797; BAA07043.1; --
DR PIR: I56258; I56258.
DR MGI: I194924; Ly78.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR Pfam: PF00560; LRR: 6.
DR Pfam: PF01463; LRRCT: 1.
DR SMART: SM00082; LRRCT: 1.
KW Receptor; Immune response; Inflammatory response; Signal; Antigen;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 661 CD180 ANTIGEN.
FT DOMAIN 21 626 EXTRACELLULAR.
FT TRANSMEM 627 650 POTENTIAL.
FT DOMAIN 651 661 CYTOPLASMIC.
FT REPEAT 51 75 LRR 1.
FT REPEAT 76 99 LRR 2.
FT REPEAT 101 123 LRR 3.
FT REPEAT 148 172 LRR 4.
FT REPEAT 174 195 LRR 5.
FT REPEAT 199 221 LRR 6.
FT REPEAT 273 296 LRR 7.
FT REPEAT 297 320 LRR 8.
FT REPEAT 322 346 LRR 9.
FT REPEAT 369 392 LRR 10.
FT REPEAT 395 418 LRR 11.
FT REPEAT 420 443 LRR 12.
FT REPEAT 444 467 LRR 13.
FT REPEAT 468 490 LRR 14.

FT REPEAT 495 518
FT REPEAT 520 540
FT REPEAT 541 565
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 22 22 T -> D (IN REF. 1; AA SEQUENCE).
FT CONFLICT 24 24 S -> N (IN REF. 1; AA SEQUENCE).
FT CONFLICT 28 28 C -> L (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 661 AA; 74266 MW; 9A04386369803C9E CRC64;

Query Match 0.8%; Score 8; DB 1; Length 661;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 NLRDLS 923
DB 371 NLRDLS 378
|||||

Search completed: July 30, 2004, 14:04:15
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:56:57 ; Search time 47 Seconds
(without alignments)
6948.114 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 1035
Sequence: 1 MURTAGDGLCLSTVLEEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 1017041 seqs, 315518202 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_invertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	740	71.5	892	Q8NEU4	Q8neU4 homo sapien
2	172	16.6	287	Q9BY26	Q9by26 homo sapien
3	12	1.2	1200	Q86W29	Q86w29 homo sapien
4	11	1.1	1043	Q86W25	Q86w25 homo sapien
5	10	1.0	501	Q8UZH1	Q8uzh1 cercopithec
6	9	0.9	509	Q9H5Z8	Q9h5z8 homo sapien
7	9	0.9	713	Q95LZ7	Q95l27 macaca fasc
8	9	0.9	846	Q8IXT0	Q8ixt0 homo sapien
9	9	0.9	1065	Q9LGI5	Q9lgis oryza sativ
10	9	0.9	1093	Q86W24	Q86w24 homo sapien
11	9	0.9	1375	Q86UB5	Q86ub5 homo sapien
12	8	0.8	43	Q8KA86	Q8ka86 chlorobium
13	8	0.8	87	Q8ZB55	Q8zb55 yersinia pe
14	8	0.8	93	Q8FR50	Q8fr50 corynebacte
15	8	0.8	123	Q9ZLP7	Q9zlp7 helicobacte
16	8	0.8	123	O25306	O25306 helicobacte

17	8	0.8	136	3	Q872H2	Q872h2 neurospora
18	8	0.8	136	16	Q70V25	Q7uv25 rhodospirell
19	8	0.8	143	2	Q50035	Q50035 mycobacteri
20	8	0.8	193	2	Q849M8	Q849m8 streptomyce
21	8	0.8	217	10	Q84S45	Q84s45 oryza sativ
22	8	0.8	237	4	Q9UOE6	Q9ue6 homo sapien
23	8	0.8	250	10	Q94L59	Q94l59 rhamnus cal
24	8	0.8	250	10	Q94L58	Q94l58 rhamnus cal
25	8	0.8	274	2	Q9F5H6	Q9f5h6 agrobacteri
26	8	0.8	274	2	Q9R725	Q9r725 agrobacteri
27	8	0.8	280	4	Q9UOE5	Q9ue5 homo sapien
28	8	0.8	284	11	Q9CV31	Q9cv31 mus musculu
29	8	0.8	300	16	Q92L81	Q92l81 rhizobium m
30	8	0.8	313	16	Q93IC6	Q93ic6 rhizobium m
31	8	0.8	317	4	Q9Y6K4	Q9y6k4 homo sapien
32	8	0.8	327	10	Q38695	Q38695 actinidia d
33	8	0.8	336	6	Q7YRA5	Q7yra5 bos taurus
34	8	0.8	338	16	Q88H07	Q88h07 pseudomonas
35	8	0.8	360	4	Q15435	Q15435 homo sapien
36	8	0.8	361	11	Q9Z105	Q9z105 mus musculu
37	8	0.8	364	16	Q8EX07	Q8ex07 mycoplasma
38	8	0.8	419	16	Q7UC70	Q7uc70 shigella fl
39	8	0.8	427	16	Q83ML0	Q83ml0 shigella fl
40	8	0.8	441	16	Q8PQP6	Q8pqp6 xanthomonas
41	8	0.8	447	4	Q96FD7	Q96fd7 homo sapien
42	8	0.8	456	11	Q924P4	Q924p4 mus musculu
43	8	0.8	456	11	Q91V17	Q91v17 mus musculu
44	8	0.8	458	9	Q7Y596	Q7y596 bacterioph
45	8	0.8	461	4	Q81ZK8	Q81zk8 homo sapien

ALIGNMENTS

RESULT 1
Q8NEU4 PRELIMINARY; PRT; 892 AA.
AC Q8NEU4; 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Monarch-1 splice form IV.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams K.L., Linhoff M.W., Ting J.P.Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV116207; AAW5145.1; -;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

Query Match 71.5%; Score 740; DB 4; Length 892;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	97	DPQETVYRVKFKRLMEDNARLGCYNLSHRYTLLLVKHSNPMVOQQQLDTRGH	156
Db	124	DPQETVYRVKFKRLMEDNARLGCYNLSHRYTLLLVKHSNPMVOQQQLDTRGH	183
QY	157	ARTVGHQASPIKIEITLFEDEERPPPTVTVMOGAGIGKSMIAHKVMDWADGKLFQGR	216
Db	184	ARTVGHQASPIKIEITLFEDEERPPPTVTVMOGAGIGKSMIAHKVMDWADGKLFQGR	243

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026393; BAB15469.1; -.
 DR HSSP; P10775; 2BNH.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_Rninh.
 DR Pfam; PF00560; LRR; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 509 AA; EB7535AF69817F5B CRC64;

Query Match 0.9%; Score 9; DB 4; Length 509;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 LLCEGLRHP 771
 |||||
 DB 207 LLCEGLRHP 215

RESULT 7
 Q95LZ7 PRELIMINARY; PRT; 713 AA.
 AC Q95LZ7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 RA Terao K., Sugano S.;
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
 RT libraries."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB071044; BAB64437.1; -.
 DR InterPro; IPR007091; LRR_Rninh.
 DR InterPro; IPR007091; NACHT_NTPase.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 713 AA; 80258 MW; 6F214C9B773F54DC CRC64;

Query Match 0.9%; Score 9; DB 6; Length 713;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 YSFIHLSFQ 491
 |||||
 DB 239 YSFIHLSFQ 247

RESULT 8
 Q8IXT0 PRELIMINARY; PRT; 846 AA.
 ID Q8IXT0
 AC Q8IXT0
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to NALP2 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC039269; AAH39269.1; -.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 SQ SEQUENCE 846 AA; 96368 MW; 46BBE8245550E39F7 CRC64;

Query Match 0.9%; Score 9; DB 4; Length 846;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 YSFIHLSFQ 491
 |||||
 DB 499 YSFIHLSFQ 507

RESULT 9
 Q9LGI5 PRELIMINARY; PRT; 1065 AA.
 ID Q9LGI5
 AC Q9LGI5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN P009G03.21 OR P0030H07.4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P009G03."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0030H07."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF002522; BAB3621.1; -.
 DR EMBL; AF003045; BAB44042.1; -.
 DR Gramine; Q9LGI5; -.
 DR GO; GO:0005224; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007090; LRR_plant.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00560; LRR; 17.
 DR Pfam; PF00669; pkinase; 1.
 DR PRINTS; PRO00019; LEURICHRPT.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1065 AA; 115226 MW; 9C221B26F64AB551 CRC64;

Query Match 0.9%; Score 9; DB 10; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 917 LRELDLSFN 925

RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AE012973; AA073495.1; -
 DR TIGR; CR2282; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 43 AA; 4973 MW; 796C897A0A19RCF0 CRC64;

Query Match 0.8%; Score 8; DB 16; Length 43;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1002 KLRVLMWF 1009
 DB 22 KLRVLMWF 29
 |||||

RESULT 13

ID Q8ZB55 PRELIMINARY; PRT; 87 AA.
 AC Q8ZB55;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE BoliA-like protein (Hypothetical protein).
 GN YPO3570 OR Y0141.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]

SEQUENCE FROM N.A.

RP STRAIN=CO-92 / Biovar Orientalis;
 RC MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RL Nature 413:523-527(2001).

RN [2]

RN SEQUENCE FROM N.A.
 RP STRAIN=KIM5 / Biovar Mediaevalis;
 RC MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM."
 RL J. Bacteriol. 184:4601-4611(2002).
 DR ENBL; AJ414157; CAC92799.1; -
 DR ENBL; AE013614; AA083735.1; -
 DR PIR; AC0434; AC0434.
 DR GO; GO:0030528; F:transcription regulator activity; IEA.
 DR InterPro; IPR002634; BOLA.
 DR Pfam; PF01722; BOLA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 87 AA; 9962 MW; 1403F68ACF7A254A CRC64;

Query Match 0.8%; Score 8; DB 16; Length 87;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 SHFQVIVV 611
 DB 29 SHFQVIVV 36
 |||||

RESULT 14

Q8FR50 PRELIMINARY; PRT; 93 AA.
 ID Q8FR50

Q8FR50;
 AC 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN CR0916.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Usuda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005217; BAC17726.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 93 AA; 10616 MW; 9B610E35B28AB75B CRC64;

Query Match 0.8%; Score 8; DB 16; Length 93;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 ELKPSFHD 278
 DB 5 ELKPSFHD 12
 |||||

RESULT 15

ID Q8ZLP7 PRELIMINARY; PRT; 123 AA.
 AC Q8ZLP7;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative flagellar motor switch protein.
 GN FLIN OR JHP0531.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]

RN SEQUENCE FROM N.A.
 RP MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori."
 RL Nature 397:176-180(1999).
 DR ENBL; AE001485; AAD06097.1; -
 DR PIR; D71922; D71922.
 DR GO; GO:0009425; C:flagellar basal body (sensu Bacteria); IEA.
 DR GO; GO:0003774; F:motor activity; IEA.
 DR GO; GO:0006935; F:chemotaxis; IEA.
 DR GO; GO:0001539; P:ciliary/flagellar motility; IEA.
 DR InterPro; IPR001172; Flagellar_FLIN.
 DR InterPro; IPR001543; Spoa.
 DR Pfam; PF01052; Spoa; 1.
 DR PRINTS; PR00956; FLGMOTORFLIN.
 DR ProDom; PD001777; Spoa; 1.
 KW Complete proteome.
 SQ SEQUENCE 123 AA; 13966 MW; 6ED16BF028BF1CC1 CRC64;

Query Match 0.8%; Score 8; DB 16; Length 123;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LSTYLEEL 20
| | | | |
Db 27 LSTYLEEL 34

Search completed: July 30, 2004, 14:03:46
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 00:15:25 ; Search time 11825 Seconds
(without alignments)
11391.979 Million cell updates/sec

Title: US-10-781-294-23
Perfect score: 3108
Sequence: 1 atgtctacgaacgcaggcag.....attggacattgctgctga 3108

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rtd.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	3108	100.0	3108	6	AX417214	AX417214 Sequence
2	3057	98.4	3300	6	AX459869	AX459869 Sequence
3	3017	97.1	3731	9	AY116204	AY116204 Homo sapi
4	3001	96.6	3507	9	BC028069	BC028069 Homo sapi
5	2997.8	96.5	3827	9	AY095146	AY095146 Homo sapi
6	2997.8	96.5	3827	9	AY154467	AY154467 Homo sapi
7	2759	88.8	3563	9	AY116205	AY116205 Homo sapi
8	2644.2	85.1	3466	9	AX833594	AX833594 Sequence
9	2644.2	85.1	3466	9	AK095460	AK095460 Homo sapi
10	2470	79.5	3395	9	AY116206	AY116206 Homo sapi
11	2417	77.8	3221	9	AY116207	AY116207 Homo sapi
12	1703	54.8	147330	9	AC008753	AC008753 Homo sapi
13	1681.8	54.1	2158	6	AX575503	AX575503 Sequence
14	1616	52.0	4931	6	AX684291	AX684291 Sequence
15	1256.2	40.4	2159	10	AY364010	AY364010 Mus muscu
16	1036.2	33.3	223344	2	AC106193	AC106193 Rattus no
17	997.8	32.1	218270	2	AC079499	AC079499 Mus muscu
18	803.6	25.9	4170	9	AF410477	AF410477 Homo sapi
19	800.4	25.8	3102	6	AX299762	AX299762 Sequence
20	800.4	25.8	3579	9	AF468522	AF468522 Homo sapi
21	800.4	25.8	3857	9	AX299760	AX299760 Sequence
22	800.4	25.8	3857	9	AF420469	AF420469 Homo sapi
23	783.4	25.2	3913	10	AF486632	AF486632 Mus muscu
24	783.4	25.2	2914	9	AY422168	AY422168 Homo sapi
25	700.2	22.5	3828	9	AF427617	AF427617 Homo sapi
26	698.6	22.5	3237	9	AY092033	AY092033 Homo sapi
27	639	20.6	1958	9	AY0511283	AY0511283 Sequence
28	639	20.6	2524	6	AX417249	AX417249 Homo sapi
29	637.4	20.5	2528	9	AF418985	AF418985 Homo sapi
30	637.4	20.5	212426	9	AC104335	AC104335 Homo sapi
31	626.6	20.2	2546	6	AX780309	AX780309 Sequence
32	625.4	20.1	2546	9	AK027194	AK027194 Homo sapi
33	595.6	19.2	2494	9	AF231021	AF231021 Homo sapi
34	589	19.0	157300	10	AL592522	AL592522 Mouse DNA
35	589	19.0	241714	2	AC083834	AC083834 Mus muscu
36	586.6	18.9	229028	2	AC123316	AC123316 Rattus no
37	586.6	18.9	229793	2	AC098184	AC098184 Rattus no
38	580.2	18.7	223791	2	AC084042	AC084042 Mus muscu
39	475.4	15.3	3282	9	AY154469	AY154469 Homo sapi
40	470.4	15.1	1557	6	AX459897	AX459897 Sequence
41	402.8	13.0	3489	6	AX478549	AX478549 Sequence
42	402.8	13.0	3830	6	AX704823	AX704823 Sequence
43	402.8	13.0	3885	9	AY154460	AY154460 Homo sapi
44	402.8	13.0	3926	6	AX704821	AX704821 Sequence
45	402.8	13.0	5859	6	AX459881	AX459881 Sequence

ALIGNMENTS

RESULT 1
AX417214
LOCUS AX417214 3108 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 23 from Patent WO0226780.
ACCESSION AX417214
VERSION AX417214.1 GI:21449784
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Read, J.C., Godzik, A., Chu, Z.L., Pawlowski, K., Fiorentino, L.,
Ariza, M.E. and Stehlik, C.
TITLE Paad domain-containing polypeptides, encoding nucleic acids, and

481	GGACACGAGCTAGCCCCATCAAGATAGAGACCCCTCTTTGAGCCAGAGGAGCGCCCC	540
541	GAGCCACCGCGCATCGTGGTGCATGCAAGCGCGGAGGGATAGCAAGTCCATGCTGGCA	600
541	GAGCCACCGCGCATCGTGGTGCATGCAAGCGCGGAGGGATAGCAAGTCCATGCTGGCA	600
601	CACAAGGTGATGCTGGACTGGCGGACGGGAAGCTCTTCCAAGCAGATTTGATTATCTC	660
601	CACAAGGTGATGCTGGACTGGCGGACGGGAAGCTCTTCCAAGCAGATTTGATTATCTC	660
661	TTCTTACATCAACTGCAGGGAGATGAACACAGAGTGCACCGGAATGCAGATGCAAGACCTC	720
661	TTCTTACATCAACTGCAGGGAGATGAACACAGAGTGCACCGGAATGCAGATGCAAGACCTC	720
721	ATCTTCAGCTGCTGGCTGAGCCACGACGGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG	780
721	ATCTTCAGCTGCTGGCTGAGCCACGACGGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG	780
781	CGGCTCTCTTTTATCATGACACGGCTTCGATGAGCTCAAGCCTTCTTTCCAGCATCTCTCAG	840
781	CGGCTCTCTTTTATCATGACACGGCTTCGATGAGCTCAAGCCTTCTTTCCAGCATCTCTCAG	840
841	GGACCTTGCTGCTCTGCTGGGAGGAGAAAAGCGCCACGGAGCTGCTTTAAACAGCTTAA	900
841	GGACCTTGCTGCTCTGCTGGGAGGAGAAAAGCGCCACGGAGCTGCTTTAAACAGCTTAA	900
901	ATTCCGAAGAAGCTGCTCCCTGAGCTATCTTTGGCTCATCCACACAGCGCCACCGGCTTG	960
901	ATTCCGAAGAAGCTGCTCCCTGAGCTATCTTTGGCTCATCCACACAGCGCCACCGGCTTG	960
961	GAGAGCTTCACCGTCTGCTGGAGACACCCAGGATGTGGAGATCTCGGGCTTCTCTGAG	1020
961	GAGAGCTTCACCGTCTGCTGGAGACACCCAGGATGTGGAGATCTCGGGCTTCTCTGAG	1020
1021	GCAGAAAGGAGGAATACCTTCTACAGATTTCCACATGACAGACGAGCGGGCCAAAGTC	1080
1021	GCAGAAAGGAGGAATACCTTCTACAGATTTCCACATGACAGACGAGCGGGCCAAAGTC	1080
1081	TTCAATTAACGTGAGGACAAACAGGCTCTCTTTCACATGTGCTTCGTCCTCGCTGTCG	1140
1081	TTCAATTAACGTGAGGACAAACAGGCTCTCTTTCACATGTGCTTCGTCCTCGCTGTCG	1140
1141	TGGTGTGTGTACTGTCCTCCAGCAGACGCTGGAGGGTGGGGGCTGTTGAGACAGACG	1200
1141	TGGTGTGTGTACTGTCCTCCAGCAGACGCTGGAGGGTGGGGGCTGTTGAGACAGACG	1200
1201	TCAGGACACCACTCAGCTGTACATGCTACTACCTGAGTCTGATGATCAACCCAAAGCG	1260
1201	TCAGGACACCACTCAGCTGTACATGCTACTACCTGAGTCTGATGATCAACCCAAAGCG	1260
1261	GGGGCCCCGGCTCCAGCCCCCACCACACAGAGAGGGTGTGCTCTTGGCGGCAGAT	1320
1261	GGGGCCCCGGCTCCAGCCCCCACCACACAGAGAGGGTGTGCTCTTGGCGGCAGAT	1320
1321	GGGCTCTGGAATCAGAAAATCCTATTTGAGGACGAGGACTCCGGGAACACGGCTTGAC	1380
1321	GGGCTCTGGAATCAGAAAATCCTATTTGAGGACGAGGACTCCGGGAACACGGCTTGAC	1380
1381	GGGGAAGACGCTCTGCTCTTCCACATGACATCTTCCAGAGGACATCAACTGTGAG	1440
1381	GGGGAAGACGCTCTGCTCTTCCACATGACATCTTCCAGAGGACATCAACTGTGAG	1440
1441	AGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAAATCTTTGACGCTATCTACTATC	1500
1441	AGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAAATCTTTGACGCTATCTACTATC	1500
1501	CTGACACGAGGGGAGCGGGGAGGCCACAGACCAAGGACGAGGCTGTTGACCGAG	1560
1501	CTGACACGAGGGGAGCGGGGAGGCCACAGACCAAGGACGAGGCTGTTGACCGAG	1560
1561	TACCGCTTTTCTGAAGAGGACTTCTGCGACCTCACAGCGCTTCTCTTTGGACTCCTG	1620
1561	TACCGCTTTTCTGAAGAGGACTTCTGCGACCTCACAGCGCTTCTCTTTGGACTCCTG	1620

QY	1621	AACGAGGACACGAGAGCCACTCGAGAGAAGAGTCTCTGCTGGAAGGTCTGCGCCGACATC	1680
DB	1621	AACGAGGACACGAGAGCCACTCGAGAGAAGTCTCTGCTGGAAGGTCTGCGCCGACATC	1680
QY	1681	AGATGGACCTGTTGCAGTGGATCCAAAGCAAAGCTCAGAGCGAGCGTCCACCCTCGAG	1740
DB	1681	AGATGGACCTGTTGCAGTGGATCCAAAGCAAAGCTCAGAGCGAGCGTCCACCCTCGAG	1740
QY	1741	CAGGCGCTCTTGGAGTCTTTCAGCTGCTGTACAGAGTCCAGGAGGAGGATTTATCCAG	1800
DB	1741	CAGGCGCTCTTGGAGTCTTTCAGCTGCTGTACAGAGTCCAGGAGGAGGATTTATCCAG	1800
QY	1801	CAGGCCCTGAGCCATTCAGGCTGATCTGCTGCAGCAACATGTCCTCCAAAGTAGGACAC	1860
DB	1801	CAGGCCCTGAGCCATTCAGGCTGATCTGCTGCAGCAACATGTCCTCCAAAGTAGGACAC	1860
QY	1861	ATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGGAGCGCCAGGTGCTGTATGGC	1920
DB	1861	ATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGGAGCGCCAGGTGCTGTATGGC	1920
QY	1921	GCCACCTACAGCGCGAGCGGGAAGACCGCGGAGGTGCTCCGAGAGCGCACGCTG	1980
DB	1921	GCCACCTACAGCGCGAGCGGGAAGACCGCGGAGGTGCTCCGAGAGCGCACGCTG	1980
QY	1981	TTGFTGCAGCTCAGACCAGAGGACCGTCTCTGTGGACGCTACAGTGAACATCTGGCA	2040
DB	1981	TTGFTGCAGCTCAGACCAGAGGACCGTCTCTGTGGACGCTACAGTGAACATCTGGCA	2040
QY	2041	CGGCGCTGTGCACCAATCCAACTGATAGAGCTGTCTGTACCGAAATGCCCTGGGC	2100
DB	2041	CGGCGCTGTGCACCAATCCAACTGATAGAGCTGTCTGTACCGAAATGCCCTGGGC	2100
QY	2101	AGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCRAACTTCAGAAC	2160
DB	2101	AGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCRAACTTCAGAAC	2160
QY	2161	CTGAGGCTGAAGAGGTCCGCACTCCAGAGCTCAGCTCGAGGACCTCTCTGAGCTCTC	2220
DB	2161	CTGAGGCTGAAGAGGTCCGCACTCTCCAGCTCAGGCTGCGAGACCTCTCTGAGCTCTC	2220
QY	2221	ATAGCCCAATAGAAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATTCAGAGC	2280
DB	2221	ATAGCCCAATAGAAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATTCAGAGC	2280
QY	2281	ATGATGCTGCTTTTCGAGGGCGCTCGCGCATCCCGAGTGGAGGTGAGTTCAGTTG	2340
DB	2281	ATGATGCTGCTTTTCGAGGGCGCTCGCGCATCCCGAGTGGAGGTGAGTTCAGTTG	2340
QY	2341	AGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTGCTCGGCAACCAAC	2400
DB	2341	AGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTGCTCGGCAACCAAC	2400
QY	2401	CCACATCTGGTTGAGTTGACCTGACAGGAATGCACTGAGAGATTTGGGCTGAGGTTA	2460
DB	2401	CCACATCTGGTTGAGTTGACCTGACAGGAATGCACTGAGAGATTTGGGCTGAGGTTA	2460
QY	2461	CTATGCCAGGCACTGAGGACCCAGCTGCGACTACGAGTCTGAGGATTTGGGCTGAGTCTGC	2520
DB	2461	CTATGCCAGGCACTGAGGACCCAGCTGCGACTACGAGTCTGAGGATTTGGGCTGAGTCTGC	2520
QY	2521	CGCCTCACTGCTGCTGTGACGAGCTGGCTCAACTCTCAGTGTGAACAGAGGCTG	2580
DB	2521	CGCCTCACTGCTGCTGTGACGAGCTGGCTCAACTCTCAGTGTGAACAGAGGCTG	2580
QY	2581	AGAGAGCTGACCTGAGCTGAATGAGTGGGGAACCTCGGGGTGCTGCTGTGTGAG	2640
DB	2581	AGAGAGCTGACCTGAGCTGAATGAGTGGGGAACCTCGGGGTGCTGCTGTGTGAG	2640
QY	2641	GGCCTCAGGCATCCACGCTGACGCTCCAGACCCCTGCGGTTGGGCACTGCGCGCTGGGC	2700
DB	2641	GGCCTCAGGCATCCACGCTGACGCTCCAGACCCCTGCGGTTGGGCACTGCGCGCTGGGC	2700

QY	2701	TCGCGCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGGAGCTG	2760
Db	2701	TCTGCGCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGGAGCTG	2760
QY	2761	GACTTGAGTTTCAACGACTCGGAGACTGGGCCTGTGGTTGCTGGCTGAGGGGCTCAA	2820
Db	2761	GACTTGAGTTTCAACGACTCGGAGACTGGGCCTGTGGTTGCTGGCTGAGGGGCTCAA	2820
QY	2821	CATCCCGCCTGCAGACTCCAGAACTGTGGCTGTGATAGCTGTGGCTCCAGACCAGAAGCT	2880
Db	2821	CATCCCGCCTGCAGACTCCAGAACTGTGGCTGTGATAGCTGTGGCTCCAGACCAGAAGCT	2880
QY	2881	TGTGAAGATCTTTACTTCACCTGGGATCAACAGACCTTTGACCGACCTTTTACTGACC	2940
Db	2881	TGTGAAGATCTTTACTTCACCTGGGATCAACAGACCTTTGACCGACCTTTTACTGACC	2940
QY	2941	AACAACGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGCAAGCGCTGAGCCATCTCTGGC	3000
Db	2941	AACAACGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGCAAGCGCTGAGCCATCTCTGGC	3000
QY	3001	TGAAAACCTCCGAGTCCCTCTGTTATTATTTGGGATGACCTGAATAAATGACCCACAGTAGG	3060
Db	3001	TGAAAACCTCCGAGTCCCTCTGTTATTATTTGGGATGACCTGAATAAATGACCCACAGTAGG	3060
QY	3061	TGGGAGCGCTTCGAGTAACAAAACCTTATTTGGACATTGGCTGCTGA	3108
Db	3061	TGGGAGCGCTTCGAGTAACAAAACCTTATTTGGACATTGGCTGCTGA	3108
<hr/>			
RESULT 2			
AX459869			
LOCUS AX459869 3300 bp DNA linear PAT 30-NOV-2002			
DEFINITION Sequence 12 from Patent WO0240668.			
ACCESSION AX459869			
VERSION AX459869.1 GI:21725643			
KEYWORDS Homo sapiens (human)			
SOURCE Homo sapiens			
ORGANISM Homo sapiens			
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE 1			
JOURNAL Techopp,J. and Martinon,F.			
PROTEINS Proteins and dna sequences underlying these proteins used for			
TREATING treating inflammations			
PATENT Patent: WO 0240668-A 12 23-MAY-2002;			
APOTECH Apotech Research and Development Ltd. (CH)			
FEATURES Location/Qualifiers			
source 1..3300			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="NALP3/PY5.dna"			
<hr/>			
ORIGIN			
Query Match 98.4%; Score 3057; DB 6; Length 3300;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCTACGAAACCGCAGCGAGGACGGCCTCTGCGCTGTCCACCTACTTGGAAAGATC	60
Db	1	ATGCTACGAAACCGCAGCGAGGACGGCCTCTGCGCTGTCCACCTACTTGGAAAGATC	60
QY	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTATACCTGGGACCGCACAGAGCTGGAGAA	120
Db	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTATACCTGGGACCGCACAGAGCTGGAGAA	120
QY	121	GGCAAGATCCCTGGGGAAGCATGAGAAGCGCGGTCCCCTGGAAATGCCCCAGCTGCTC	180
Db	121	GGCAAGATCCCTGGGGAAGCATGAGAAGCGCGGTCCCCTGGAAATGCCCCAGCTGCTC	180
QY	181	ATCACCACCTTCGGGCCAGAGAGCGCTTGGAGTTGGCTCTCAGCACTTTGACGGATA	240
Db	181	ATCACCACCTTCGGGCCAGAGAGCGCTTGGAGTTGGCTCTCAGCACTTTGACGGATA	240


```
Db 2401 CCATCTCTGTTGAGTTGGACCTGACAGGAATGCACTGGAGGATTTGGGCTGAGGTTA 2460
QY 2461 CTATGACAGGACTGAGGACCCAGCTGTCAGACTACGACTTTGGCTGGAAGATCTGC 2520
Db 2461 CTATGACAGGACTGAGGACCCAGCTGTCAGACTACGACTTTGGCTGGAAGATCTGC 2520
QY 2521 CGCCTCACTGCTGCTGCTGTCAGAGCTGGCCCTCAACTCTCAGTGTGAACAGAGCCTG 2580
Db 2521 CGCCTCACTGCTGCTGCTGTCAGAGCTGGCCCTCAACTCTCAGTGTGAACAGAGCCTG 2580
QY 2581 AGAGAGCTGACCTGAGCTGGAATGAGCTGGGGGAACTCGGGGTGCTGCTGTGTGAG 2640
Db 2581 AGAGAGCTGACCTGAGCTGGAATGAGCTGGGGGAACTCGGGGTGCTGCTGTGTGAG 2640
QY 2641 GGCCTCAGGACATCCACGCTGCAAGCTCCAGACCTCGGCTGGGCTGAGCTGCGGGCTGGC 2700
Db 2641 GGCCTCAGGACATCCACGCTGCAAGCTCCAGACCTCGGCTGGGCTGAGCTGCGGGCTGGC 2700
QY 2701 TCTGCGCCCTGTGAGGCTCTTTCTGTGCTGCTCCAGGCCAACCAACCACTCCCGGAGCTG 2760
Db 2701 TCTGCGCCCTGTGAGGCTCTTTCTGTGCTGCTCCAGGCCAACCAACCACTCCCGGAGCTG 2760
QY 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGCTGCTGAGGGCTGCAA 2820
Db 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGCTGCTGAGGGCTGCAA 2820
QY 2821 CATCCGCTGACAGCTCCAGAACTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATGCTG 2880
Db 2821 CATCCGCTGACAGCTCCAGAACTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATGCTG 2880
QY 2881 TGTGAGAACTTTACTTCACTCCCTGGGATCAACGACACTTGACGACCTTTACCTGACC 2940
Db 2881 TGTGAGAACTTTACTTCACTCCCTGGGATCAACGACACTTGACGACCTTTACCTGACC 2940
QY 2941 AACACGCTGAGGACACAGCTGTCAGCTGCTTTGCAAGCGCTGAGCCATCCTGGC 3000
Db 2941 AACACGCTGAGGACACAGCTGTCAGCTGCTTTGCAAGCGCTGAGCCATCCTGGC 3000
QY 3001 TGCACAACTCCGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3057
Db 3001 TGCACAACTCCGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3057

RESULT 3
AY116204
LOCUS AY116204 3731 bp mRNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens monarch-1 mRNA, complete cds; alternatively spliced.
ACCESSION AY116204
VERSION AY116204.1 GI:21711820
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3731)
AUTHORS Williams,K.L., Taxman,D.J., Linhoff,M.W., Reed,W. and Ting,J.P.Y.
TITLE Monarch-1: A Pylrin/Nucleotide-Binding Domain/Leucine-Rich Repeat
Protein That Controls Classical and Nonclassical MHC Class I Genes
JOURNAL J. Immunol. 170 (11), 5354-5358 (2003)
PUBMED 12759408
REFERENCE 2 (bases 1 to 3731)
AUTHORS Williams,K.L., Linhoff,M.W., Harton,J.A. and Ting,J.P.Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA
FEATURES
Location/Qualifiers
1..3731
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
```

```
/cell_line="U937"
221..3409
/notes="alternatively spliced"
/codon_start=1
/product="monarch-1"
/protein_id="AAM75142.1"
/db_xref="GI:21711821"
/translation="MLRTAGRDLGCLRLSYLBLELAVELKKPKLYLGTATELGKIP
WSMEKAGLEMAQLLTHFPEEAWRLALSTFERINRKLMDWEGREDLVRDTPPG
PSLGNQSCLLLEVLDTGRGHARTVYRRKFRMLMEDNRALGECVNLSHRYRL
LIVKHSNPMQVQQLDTRGHARTVYRRKFRMLMEDNRALGECVNLSHRYRL
AGTGLSKLHAKVWLMDWADGLQGRDFYLYINCRMNOSATECSMODLIFSCWPPPS
APQELIRVPERLLFTIDGDELKPSFHPDQGPWCLCWEKRPTELNSLIRKLLP
ELSLITPTALEKLHLEHPRHVEILGFSEAEKEYFYKPHNABAGQVFNVR
DNEPLTFMCFELVWVCTCQQLEGGLLRQTSRTTAVMLVLLSLMDPKPAP
RHOFPNQGLCLAAQWLNQKILFEEQDLRRKGLDGEDVSAFLNMNIFQKINCEP
YYSFTHLSQEEFAAMYIIDEGBGAGPDQDVTRLLTYAFSEFSLALTSLRFLFGL
LNEETRSHLEKSLCWKSPHIMDLLOWIQSKAQSDGSLQOQSLFFSCLVEIOEE
FIQQLSHFQVIVWSNIAKMBHVSFCLKCRSAQVHLHYGATYSADGEARCSA
GAHTLLVQLRPERTVILLDAYSEHLAAALCTNPNLIELSYRNALSGRVKLLCQGRH
PNCKLQNRKKCRISSSACEDLSAALIANKNLRLMDLSNGVGFPMWLLCEGLRHP
CRLQMIQRKQLESAGCAEMASVLGTNPHLVELDLTGNALEDGLRLCCGLRHP
CRLQMIQRKQLESAGCAEMASVLGTNPHLVELDLTGNALEDGLRLCCGLRHP
KLQTLRLGICRLGSAACEGLSVLQANHLRLDLSPNDLGDWGLMLLAEGLOHPCR
LQKLWLDSCGLTAKACENLYFTLGINOTLDTLYLTNNALGDTGVRLCLKRLSHPGCKL
RVMLFGMDLNKMTSHRLAALRVTPYLDIGC"
859..892
/notes="Region: P-loop"
2069..3406
/notes="Region: leucine rich repeat"
221..289
/notes="Region: pyrin domain"

misc_feature
859..892
/notes="Region: P-loop"
2069..3406
/notes="Region: leucine rich repeat"
221..289
/notes="Region: pyrin domain"

ORIGIN
Query Match 97.1%; Score 3017; DB 9; Length 3731;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 81; Gaps 1;
QY 1 ATGCTACGAACCGCAGGAGGAGCGCCCTCTGTGGCTGTCTCCACCTTCTGGAAGAACTC 60
Db 221 ATGCTACGAACCGCAGGAGGAGCGCCCTCTGTGGCTGTCTCCACCTTCTGGAAGAACTC 280
QY 61 GAGGCTGTGGAAGTCAAGTTATACCTGGGACCGCGACAGAGCTGGGAGAA 120
Db 281 GAGGCTGTGGAAGTCAAGTTATACCTGGGACCGCGACAGAGCTGGGAGAA 340
QY 121 GCAAGATCCCTGGGAGAGCATGAGAGGCGGCTCCCTCGAAATGGCCAGCTGCTC 180
Db 341 GCAAGATCCCTGGGAGAGCATGAGAGGCGGCTCCCTCGAAATGGCCAGCTGCTC 400
QY 181 ATCAGCCACTTGGGCGCAGAGGCGCTGGAGTTGGCTCTCAGCAGCTTTGAGCGGATA 240
Db 401 ATCAGCCACTTGGGCGCAGAGGCGCTGGAGTTGGCTCTCAGCAGCTTTGAGCGGATA 460
QY 241 AACAGAGAGGACTCTGGGAGAGAGGACAGAGAGGAGGACTGTGTGAGG----- 288
Db 461 AACAGAGAGGACTCTGGGAGAGAGGACAGAGAGGAGGACTGTGTGAGGAGATACCCACT 520
QY 289 ----- 288
Db 521 GGTGGCCGCTCTCACTTTGGGAACCACTCAATGCTCTTGGAGTCTCTCTTGTCACT 580
QY 289 -----GATCCCGAGAAACCTTACAGGAGCTATGTCGGCAGAGAAATTCGGGCTCATG 339
Db 581 CCAAGAAAGATCCCGAGAAACCTTACAGGAGCTATGTCGGCAGAGAAATTCGGGCTCATG 640
QY 340 GAAGACCGCAATCGCGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399
Db 641 GAAGACCGCAATCGCGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 700
QY 400 CTGCTGGTGAAGAGGAGCACTCAAAACCCCATGAGGTCAGAGCAGCAGCTTCTTGACACAGC 459
Db 701 CTGCTGGTGAAGAGGAGCACTCAAAACCCCATGAGGTCAGAGCAGCAGCTTCTTGACACAGC 760
```


Db	1860	TGGAAGGTCTCGCCGCAATCAAGATGGAAGCTGTTCAGTGGATCCAAAGCAAGACTCAG	1911
Qy	1720	AGCGACGGCTCCACCTCTGCAGCAGGGCTCCTTTGGAGTTCTTTCAGCTCTGTGTACGAGATC	1779
Db	1920	AGCGAAGGCTCCACCTCTGCAGCAGGGCTCCTTTGGAGTTCTTTCAGCTCTGTGTACGAGATC	1979
Qy	1780	CAGGAGGAGGATTTATCCAGCAGGCCTTGAGCCACTTCCAGGTGATCGTGGTCAAGCAAC	1839
Db	1980	CAGGAGGAGGATTTATCCAGCAGGCCTTGAGCCACTTCCAGGTGATCGTGGTCAAGCAAC	2039
Qy	1840	ATTGCCCTCCAAAGATGGAGCAGATGGTCTCCTCGTTCTGTCTGAAGCGCTGCAGAGGCGCC	1899
Db	2040	ATTGCCCTCCAAAGATGGAGCAGATGGTCTCCTCGTTCTGTCTGAAGCGCTGCAGAGGCGCC	2099
Qy	1900	CAGGTGCTGCACCTTGATGGCGCCACCTACAGCGCGGACCGGGAAGACCGCGCGAGGTGC	1959
Db	2100	CAGGTGCTGCACCTTGATGGCGCCACCTACAGCGCGGACCGGGAAGACCGCGCGAGGTGC	2159
Qy	1960	TCCGAGGAGGCGACACGCTTGTTGGTGCAGCTCAGACACAGAGAGGACCGTTCCTGCTGGAC	2019
Db	2160	TCCGAGGAGGCGACACGCTTGTTGGTGCAGCTCAGACACAGAGAGGACCGTTCCTGCTGGAC	2216
Qy	2020	GCCTACAGTCAACATCTGGCAGCGGCGCTGTGCACCAATCCAAACCTGATAGAGCTGTCT	2079
Db	2217	GCCTACAGTCAACATCTGGCAGCGGCGCTGTGCACCAATCCAAACCTGATAGAGCTGTCT	2276
Qy	2080	CTGTACCGAAATGCCCTTGGGCAGCGGGGGGTGAAGCTGCTCTGTCAAGAGCTCAGACAC	2139
Db	2277	CTGTACCGAAATGCCCTTGGGCAGCGGGGGGTGAAGCTGCTCTGTCAAGAGCTCAGACAC	2336
Qy	2140	CCCAACTGCAAACTTCAGAACTTGAGGCTGAAGAGGTGCGCAATCTCCAGCTCAGGCTGC	2199
Db	2337	CCCAACTGCAAACTTCAGAACTTGAGGCTGAAGAGGTGCGCAATCTCCAGCTCAGGCTGC	2396
Qy	2200	GAGGACCTCTCTCGAGCTCTCATAGCCATAAGAAATTTGACAGAGATGATCTCATGTGGC	2259
Db	2397	GAGGACCTCTCTCGAGCTCTCATAGCCATAAGAAATTTGACAGAGATGATCTCATGTGGC	2456
Qy	2260	AACGGGCTTGGATTCACAGGCATGATGCTGCTTTGCGAGGGCTCGGGCATCCCCAGTGC	2319
Db	2457	AACGGGCTTGGATTCACAGGCATGATGCTGCTTTGCGAGGGCTCGGGCATCCCCAGTGC	2516
Qy	2320	AGGCTCAGATGATTCAGTTGAGGAAGTGCAGCTGGAGTTCGGGGCTTGTCAAGGAGATG	2379
Db	2517	AGGCTCAGATGATTCAGTTGAGGAAGTGCAGCTGGAGTTCGGGGCTTGTCAAGGAGATG	2576
Qy	2380	GCCTTCGTGCTCGGCACCAACCCACATCTGGTTGAGTTTGGACCTGACAGAAATGCACGT	2439
Db	2577	GCCTTCGTGCTCGGCACCAACCCACATCTGGTTGAGTTTGGACCTGACAGAAATGCACGT	2636
Qy	2440	GAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGCAACCCAGTCTGACAGCTACGG	2499
Db	2637	GAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGCAACCCAGTCTGACAGCTACGG	2696
Qy	2500	ACTTTGTGCTGAAGATTCGCGCCTCACTGTGTGCTGCTGTGACAGAGCTGGCCCTCAACT	2559
Db	2697	ACTTTGTGCTGAAGATTCGCGCCTCACTGTGTGCTGCTGTGACAGAGCTGGCCCTCAACT	2756
Qy	2560	CTCAGTGTCAACACAGAGCTGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGACCTC	2619
Db	2757	CTCAGTGTGAACACAGAGCTGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGACCTC	2816
Qy	2620	GGGGTGTCTGTGTGTGAGGGCTTCAGGCAATCCCAAGTCCAGACCTCCAGACCTTCGGG	2679
Db	2817	GGGGTGTCTGTGTGTGAGGGCTTCAGGCAATCCCAAGTCCAGACCTTCAGACCTTCGGG	2876
Qy	2680	TTGGGGCATCTGCGGCTGGGCTCTGCGCCTGTGAGGGTCTTTCTGTGTGTCTCCAGGCC	2739
Db	2877	TTGGGGCATCTGCGGCTGGGCTCTGCGCCTGTGAGGGTCTTTCTGTGTGTCTCCAGGCC	2936
Qy	2740	AACCAACACTCGGGAGCTGACCTTTCAGGACCTGGGAGACTGGGGCTCTGTGG	2799
Db	2937	AACCAACACTCGGGAGCTGACCTTTCAGGACCTGGGAGACTGGGGCTCTGTGG	2996

QY	2800	TTGCTGGCTGAGGGGCTGCACAAATCCCGCCTGCAGACTCCAGAACTGTGGCTGGATAGC	2851
Db	2997	TTGCTGGCTGAGGGGCTGCAACATCCCGCCTGCAGACTCCAGAACTGTGGCTGGATAGC	3056
QY	2860	TGTGGCCTCACAGCCCAAGGCTTGTGAAATCTTTTACCTACCCCTGGGGATCAACAGACC	2919
Db	3057	TGTGGCCTCACAGCCCAAGGCTTGTGAAATCTTTTACCTACCCCTGGGGATCAACAGACC	3116
QY	2920	TTGACCCACCTTTACCTGACCAACAACCGCCCTAGGGGACACAGGTGTCGACCTGCTTTGC	2979
Db	3117	TTGACCCACCTTTACCTGACCAACAACCGCCCTAGGGGACACAGGTGTCGACCTGCTTTGC	3176
QY	2980	AAGCGGCTGAGCCATCTCGCTGCAAACTCCGAGTCTCTGGTATTATTTGGATGGACCTG	3039
Db	3177	AAGCGGCTGAGCCATCTCGCTGCAAACTCCGAGTCTCTGGTATTATTTGGATGGACCTG	3236
QY	3040	AATAAATGACCCACAGTAGTGGGAGCGCTTCGAGTAACAAACCTTATTTGGACATT	3099
Db	3237	AATAAATGACCCACAGTAGTGGGAGCGCTTCGAGTAACAAACCTTATTTGGACATT	3296
QY	3100	GGCTGCTGA 3108	
Db	3297	GGCTGCTGA 3305	
RESULT 5			
AY095146			
LOCUS			
DEFINITION	Homo sapiens PYRIN-containing APAFI-like protein 7 mRNA	linear	PRI 12-AUG-2002
ACCESSION	AY095146		
VERSION	AY095146.1	GI:21314906	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

/translations="MLRTGFGDGLCSLYLEEAVALCKPKLYLGYATLEGGKJL
WGSMEXAGLEMAQLLTHGFPSEAWRLSTFENRKLWGERGQEDVJLPPG
PSSJLNGQSLCLLEVSJLWTRKPOQETRYVYRKERFLMEDNRALGECVNLSHYATL
LVKHSJNPMQVQQLLDGRGHARTVGHQAQSKIETLFPEDRPEPPRYVMQA
AGTCKSLMHLWMLDADGKLFQGRGDFLYFINCREMCAQSTBCSMQDLISCPWEP
APQELIRYAPRLLVLDIFGDSGLKPSFDHPQGMCKEERKPTLNLSTLRKLLP
ELSLITTPRPLLEKLHRLHPRHVEILGFSAEARKYFYFYFHNAQAGQGVNVR
DNEPLFTTFCVPLVCVWVTCLOOEGGLLROTSTKTTAVVWMLLSLQMPKRG

RLQPPNQRGLCSLAADGLWNKILFEEDQRLKXHLGDEVSFLNMMNQDKINDER
YYSFHLSPFPAAMVYILDEGEAGAGDPDVRLLITEYAFSEFFSLYTRFLFGL
LNETSHLSCKWVSPHILKMDLQWQKQSDGSLTQOQSFSLYTRFLFGL
FIQALSHFOVIVVSNIAKMEHWSFCLKCRSAQVHLXGATYSADGEDRARC
GATHLVQDPERIVLDAISEHAAALCTNPNLIELSYRNALSGRVKLLCOGLRHP
NCKLQRLKRCRISACEDLSAALIANKNLMDLSGNGVFPGMMLCEBGLRHP
CLQMTQLRKCQLESACQEMASVLTGNPHLVELDTGNALEDGLRLCLCOGLRHP
RLRLTLKICRLTAACDELASTLSVNSQSLRELDLSNELGDLGLVLLCEGLRHP
LOTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLWLLAEGLHPCRL
OKLWLSOGLTAKACENLYETLGINOTLTLXLTNNALGDTGVRLILCKRLSHFPGCKLR
VLWLFMDLNNMTHSRLLAALRVTKPYLDIGC"

ORIGIN

Query Match	96.5%;	Score 2997.8;	DB 9;	Length 3827;
Best Local Similarity	97.3%;	Pred. No. 0;		
Matches 3103;	Conservative 0;	Mismatches 2;	Indels 84;	Gaps 2;
QY	1	ATGCTACGAACCGCAGCAGGACGGCCCTCTGTGCGCTCTGTCCACCTTACTTGAAGAACA	60	
DB	177	ATGCTACGAACCGCAGCAGGACGGCCCTCTGTGCGCTCTGTCCACCTTACTTGAAGAACA	236	
QY	61	GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGAGAA	120	
DB	237	GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGAGAA	296	
QY	121	GGCAAGATCCCTGGGGAAGCATGGAAGCCGGTCCCTGGAATGCCCCAGCTGCTC	180	
DB	297	GGCAAGATCCCTGGGGAAGCATGGAAGCCGGTCCCTGGAATGCCCCAGCTGCTC	356	
QY	181	ATCACCCACTTCGGGCGCAGAGGAGCTGTGAGTTGGTCTCTAGACCTTTGAGCGGATA	240	
DB	357	ATCACCCACTTCGGGCGCAGAGGAGCTGTGAGTTGGTCTCTAGACCTTTGAGCGGATA	416	
QY	241	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGG	288	
DB	417	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGGATACCCACCT	476	
QY	289	-----	288	
DB	477	GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCTCTTGTGAAGTCTCTCTTGTCACT	536	
QY	289	-----GATCCCGAGGAACCTACAGGACATATGTCCGCGAGGAATTCGGGCTCATG	339	
DB	537	CCAAAGAAAGATCCCGAGGAACCTACAGGACATATGTCCGCGAGGAATTCGGGCTCATG	596	
QY	340	GAAGACCGCAATCGCGCTTAGGGGAATGTCTCAACCTCAGCCCGGTACACCGGCTC	399	
DB	597	GAAGACCGCAATCGCGCTTAGGGGAATGTCTCAACCTCAGCCCGGTACACCGGCTC	656	
QY	400	CTGCTGGTGAAGAGACACTCAAAACCCATCAGGTTCCAGCAGCAGCTTCTGGACACAGGC	459	
DB	657	CTGCTGGTGAAGAGACACTCAAAACCCATCAGGTTCCAGCAGCAGCTTCTGGACACAGGC	716	
QY	460	CGGGGACACGCGAGGACCGTGGGACACACGAGCTAGCCCATCAAGATAGAGACCTCTTT	519	
DB	717	CGGGGACACGCGAGGACCGTGGGACACACGAGCTAGCCCATCAAGATAGAGACCTCTTT	776	
QY	520	GAGCCAGACAGGAGCGCCCCGAGCCACCGCGACCGTGTGATCTGAAGCGCGGACAGG	579	
DB	777	GAGCCAGACAGGAGCGCCCCGAGCCACCGCGACCGTGTGATCTGAAGCGCGGACAGG	836	
QY	580	ATAGGCAAGTCCATGCTGGACACAAAGGTGATGCTGGAGCTGGCGGACGGGAAGCTCTTC	639	
DB	837	ATAGGCAAGTCCATGCTGGACACAAAGGTGATGCTGGAGCTGGCGGACGGGAAGCTCTTC	896	
QY	640	CAAGGCAAGTTTGATTTATCTCTTCTCATCACTGACGGGAGATGAACACAGAGTGCCACG	699	
DB	897	CAAGGCAAGTTTGATTTATCTCTTCTCATCACTGACGGGAGATGAACACAGAGTGCCACG	956	
QY	700	GAATGACAGCATGCAAGACCTCATCTTCACTGCTGGCTGAGCCCGCGCTCTCCAG	759	
DB	957	GAATGACAGCATGCAAGACCTCATCTTCACTGCTGGCTGAGCCCGCGCTCTCCAG	1016	

QY	760	GAGCTCATCCGAGTTCCGAGCGCCTCTTTTTCATCATCAACGGCTTCGATGAGCTCAAG	819
DB	1017	GAGCTCATCCGAGTTCCGAGCGCCTCTTTTTCATCATCAACGGCTTCGATGAGCTCAAG	1076
QY	820	CTTCTTTTCCACCATCTCAGGAGCCTCTGCTGCTGCTGGAGGAGAAACGCCCCACG	879
DB	1077	CTTCTTTTCCACCATCTCAGGAGCCTCTGCTGCTGCTGGAGGAGAAACGCCCCACG	1136
QY	880	GAGCTGCTCTTTAAACAGCTTAATTCGGAAGAAGCTGCTCCTGAGCTATCTTTGCTCATC	939
DB	1137	GAGCTGCTCTTTAAACAGCTTAATTCGGAAGAAGCTGCTCCTGAGCTATCTTTGCTCATC	1196
QY	940	ACCACACGGCCCGCTTTGGAGAGCTCCACGCTGCTGCTGGAGCAACCCAGCATGTG	999
DB	1197	ACCACACGGCCCGCTTTGGAGAGCTCCACGCTGCTGCTGGAGCAACCCAGCATGTG	1256
QY	1000	GAGATCCTGGGCTTCTCTGAGGAGAGAAAGAAAGTACTTCTACAAAGTATTTTCCAAAT	1059
DB	1257	GAGATCCTGGGCTTCTCTGAGGAGAGAAAGAAAGTACTTCTACAAAGTATTTTCCAAAT	1316
QY	1060	GCAGAGCAGCGCGGCCCAAGTCTTCAATTACGTGAGGAGCAACAGAGCTCTTCTTCCCATG	1119
DB	1317	GCAGAGCAGCGCGGCCCAAGTCTTCAATTACGTGAGGAGCAACAGAGCTCTTCTTCCCATG	1376
QY	1120	TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1179
DB	1377	TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1436
QY	1180	GGGGGGCTGTGTGACAGACAGCTCCAGGACACCACTGAGTGTACATGCTCTGCTGCTG	1239
DB	1437	GGGGGGCTGTGTGACAGACAGCTCCAGGACACCACTGAGTGTACATGCTCTGCTGCTG	1496
QY	1240	AGTCTGTATGCAACCAACCGCGGGCCCGCGCTCCAGGCGCCCAACCAACAGAGGG	1299
DB	1497	AGTCTGTATGCAACCAACCGCGGGCCCGCGCTCCAGGCGCCCAACCAACAGAGGG	1556
QY	1300	TTGTGCTCTTGGCGGAGATGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1359
DB	1557	TTGTGCTCTTGGCGGAGATGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1616
QY	1360	CTCCGGAAGCAGCGCTTAGACGGGAGAGCGCTCTGCTGCTTCTTCTCAACATGAAATCTTC	1419
DB	1617	CTCCGGAAGCAGCGCTTAGACGGGAGAGCGCTCTGCTGCTTCTTCTCAACATGAAATCTTC	1676
QY	1420	CAGAAGACATCAACTGTGAGAGGTACTACAGTCTTCACTTCACTTCACTTCACTTCACTT	1479
DB	1677	CAGAAGACATCAACTGTGAGAGGTACTACAGTCTTCACTTCACTTCACTTCACTTCACTT	1736
QY	1480	TTTTCAGCTATGTACTATATCTCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1539
DB	1737	TTTTCAGCTATGTACTATATCTCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1796
QY	1540	GTGACACGCTGTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCTTGGCTTCAACAGC	1599
DB	1797	GTGACACGCTGTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCTTGGCTTCAACAGC	1856
QY	1600	CGCTTCTTGTGTGAGTCTCTGAAACAGGAGACCGAGGCGGCGGCGGCGGCGGCGGCGG	1659
DB	1857	CGCTTCTTGTGTGAGTCTCTGAAACAGGAGACCGAGGCGGCGGCGGCGGCGGCGGCGG	1916
QY	1660	TGGAGGTCTCGCGCATCAAGATGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	1719
DB	1917	TGGAGGTCTCGCGCATCAAGATGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	1976
QY	1720	AGCGACGCTCCACCTGCGAGCGGCTCTTGGAGTCTTCTTGGAGTCTTCTTGGAGTCTTCT	1779
DB	1977	AGCGACGCTCCACCTGCGAGCGGCTCTTGGAGTCTTCTTGGAGTCTTCTTGGAGTCTTCT	2036
QY	1780	CAGGAGGAGGAGTGTATCCAGAGGCGCTGAGCCTTCCAGGTCATCTGTGTGAGTGTGAGT	1839
DB	2037	CAGGAGGAGGAGTGTATCCAGAGGCGCTGAGCCTTCCAGGTCATCTGTGTGAGTGTGAGT	2096
QY	1840	ATTGCTCCAAAGATGAGACATGCTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1899

Db 2097 ATTGCTCCAAAGATGAGACATGGTCTCTCTGTTCTGTGCTGAAGCGCTCGAGGAGCGCC 2156
QY CAGGTGCTGACCTTATGCGCCACCTACAGCGCGGACGGGGAAGACCGCGGAGGTGC 1959
Db 2157 CAGGTGCTGACCTTATGCGCCACCTACAGCGCGGACGGGGAAGACCGCGGAGGTGC 2216
QY TCCGAGGAGCGCACACGCTGTGTGTGTCAGCTCAGACACAGAGAGGACCGTTCCTGCTGGAC 2019
Db 2217 TCCGAGGAGCGCACACGCTGTGTGTGTCAGCTCAGACACAGAGAGGACCGTTCCTGCTGGAC 2273
QY GCTACAGTGAACATCTGCGAGGCGCCCTGTGCA CCAATCCAAACCTGATAGAGTGTCT 2079
Db 2274 GCTACAGTGAACATCTGCGAGGCGCCCTGTGCA CCAATCCAAACCTGATAGAGTGTCT 2333
QY CTGTACCGAAATGCCCTGCGGAGCGGGGGGTGAAGCTCTCTGTCAAGGACTCGACAC 2139
Db 2334 CTGTACCGAAATGCCCTGCGGAGCGGGGGGTGAAGCTCTCTGTCAAGGACTCGACAC 2393
QY CCAACTGCAAACTTCAAGAACCTGAGGTGAAGAGGTGCGCATCTCCAGCTCAGCTGC 2199
Db 2394 CCAACTGCAAACTTCAAGAACCTGAGGTGAAGAGGTGCGCATCTCCAGCTCAGCTGC 2453
QY GAGGACCTCTGCGAGCTCTCATAGCCATAGAAATTTGACAAGGATGGATCTCAGTGGC 2259
Db 2454 GAGGACCTCTGCGAGCTCTCATAGCCATAGAAATTTGACAAGGATGGATCTCAGTGGC 2513
QY AACCGCGTTGGATTCGCCAGCATGCTCTTTCGAGGGCTCGGGCATCCCCAATGC 2319
Db 2514 AACCGCGTTGGATTCGCCAGCATGCTCTTTCGAGGGCTCGGGCATCCCCAATGC 2573
QY AGGCTGCAAGTATTCAGTTGAGGAGTGTCAAGTGGAGTCCGGGCTTGTCAAGGATG 2379
Db 2574 AGGCTGCAAGTATTCAGTTGAGGAGTGTCAAGTGGAGTCCGGGCTTGTCAAGGATG 2633
QY GCTTCTGTGCTCGGACCAACCACTGCTGTGTGAGTGGACCTGACAGGAATGCACTG 2439
Db 2634 GCTTCTGTGCTCGGACCAACCACTGCTGTGTGAGTGGACCTGACAGGAATGCACTG 2693
QY GAGGATTTGGGCTGAGGTACTATGTCAGGAGCTGAGGACCCAGTCTGAGAGTACGG 2499
Db 2694 GAGGATTTGGGCTGAGGTACTATGTCAGGAGCTGAGGACCCAGTCTGAGAGTACGG 2753
QY ACTTTGGTGGCTGAAGATTCGCGCTCACTGCTGCTGTGAGAGCTGGCCCTCAACT 2559
Db 2754 ACTTTGGTGGCTGAAGATTCGCGCTCACTGCTGCTGTGAGAGCTGGCCCTCAACT 2813
QY CTCAGTGTGAACAGAGCTGAGAGAGCTGAGACCTGAGCTGATGAGCTGGGGACCTC 2619
Db 2814 CTCAGTGTGAACAGAGCTGAGAGAGCTGAGACCTGAGCTGATGAGCTGGGGACCTC 2873
QY GGGGTGCTGCTGCTGTGAGGGCTCAGGCATCCAGTGTCAAGCTCCAGACCTCGCGG 2679
Db 2874 GGGGTGCTGCTGCTGTGAGGGCTCAGGCATCCAGTGTCAAGCTCCAGACCTCGCGG 2933
QY TTGGGATCTCCGGCTGGGCTCTGCGCCCTGTGAGGGTCTTCTGTGCTGCTCAGGCC 2739
Db 2934 TTGGGATCTCCGGCTGGGCTCTGCGCCCTGTGAGGGTCTTCTGTGCTGCTCAGGCC 2993
QY AACCAACCTCCGGAGCTGAGTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
Db 2994 AACCAACCTCCGGAGCTGAGTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 3053
QY TTGCTGGCTGAGGGCTGCAACATCCCGCTGAGACTCCAGAACTGTGCTGGATGAC 2859
Db 3054 TTGCTGGCTGAGGGCTGCAACATCCCGCTGAGACTCCAGAACTGTGCTGGATGAC 3113
QY TGTGGCTTCAAGCAAGGCTTGTGAGAACTTTTACTTCACTTGGGATCAACAGAC 2919
Db 3114 TGTGGCTTCAAGCAAGGCTTGTGAGAACTTTTACTTCACTTGGGATCAACAGAC 3173
QY TTGACCGACCTTTACTGACCAACAAAGCCCTAGGGACACAGGTGTCGACTGCTTGC 2979

Db 3174 TTGACCGACCTTTACTGACCAACAAAGCCCTAGGGACACAGGTGTCCAGCTGCTTTC 3233
QY AAGCGCTGAGCATCTCTGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACCTG 3039
Db 3234 AAGCGCTGAGCATCTCTGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACCTG 3293
QY AATAAATGACCCACACAGTGTGGCAGCGCTTCGAGTAAACAAACCTTATTTGGACATT 3099
Db 3294 AATAAATGACCCACACAGTGTGGCAGCGCTTCGAGTAAACAAACCTTATTTGGACATT 3353
QY GGCTGCTGA 3108
Db GGCTGCTGA 3362

AY154467 3827 bp mRNA linear PRI 20-FEB-2003
Homo sapiens NALP12 (NALP12) mRNA, complete cds.
AY154467
AY154467.1 GI:28436377

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3827)
Tschoopp, J., Martinon, F. and Burns, K.
NALP12: a novel protein family involved in inflammation
Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
22451042
PUBMED 12563287

2 (bases 1 to 3827)
Martinon, F., Hofmann, K. and Tschoopp, J.
Direct Submission
Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
Location/Qualifiers
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3827
/gene="NALP12"
177..3362
/gene="NALP12"
/note="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP12"
/protein_id="AAO18163.1"
/db_xref="GI:28436378"

WGSMEKAGPMEALIIITHFGEPAWRIALSTFERINRKLWEGOREDLVDDTPPG
PSSIGNOSTCLLEVSATPRKPOETVYRKRFLMEORNAELGCVNLSHYTRL
LIVKHSNPQVQQLDGTGRHARTVGHQASPIKIELTFEEDERPEPTVWQGA
AGIKSMIAKVMWDWADKLFQGRFDYLFYINCRMNQSAECMDLIFSCWPEPS
ELSLITRTALEKHLRLEHPRHVEILGFSAEARKEYFYKPHNAEQAGOVNVR
DNEPLFTWCVPVLVWVCTCLOQLEGGLRQTSRTTAVYMLYLLSLMQPGAP
RLOPPNORGLCSLAADGLNOKILFEQDLRKHGLDGEDVSAFLNMNIFOKDINCR
YYSFHLSPFEPAAMYIIDEDEGGAGPQDVTLLTETAFASFLATSLRFLGL
LNRETRSHLSKLCWVSPHIMDLQWIKRQASQSGSTLQQSLEFSCLEYQEBE
FIQOALSHFQVIVSNATKMEHVSFCIKRQSAQLVHLGYTADGEDRARCBA
GATLIVQLPERTVLLDAYSEHLAALCTNPNIIELSLYNALSGRVKLLCOGLRHP
NCKLONLRKCRITSSACEDLSAALTANKNLTMDLSGNGVPGPGMWLLCEGLRHPQ
CRQMQLRCKQESGACQEMASVLGTNPHVELDITGNALDELGLRLCOGLRHPVC
RLTLMKICRLTAACDEASLTVSNQSLRELDLSLNEGLDGLVLLAEGLOHPACRL
LQTLRIGICRGSAAECGLSVLOAHNRELDSFNDLGDWGLWLLAEGLOHPACRL
QKWLWDSGGLTAKACENLYFTLQNTLITDLYLNNALGDTGVRLKCKLSHPGCKLR
VLWLFGLMDLNKMTSHSLAALRVTKPYLDIGC"

RESULT 6
AY154467
LOCUS
DEFINITION
Homo sapiens NALP12 (NALP12) mRNA, complete cds.
ACCESSION
AY154467
VERSION
AY154467.1
KEYWORDS
GI:28436377
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 3827)
Tschoopp, J., Martinon, F. and Burns, K.
NALP12: a novel protein family involved in inflammation
Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
22451042
PUBMED 12563287

REFERENCE
2 (bases 1 to 3827)
Martinon, F., Hofmann, K. and Tschoopp, J.
Direct Submission
Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
Location/Qualifiers
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3827
/gene="NALP12"
177..3362
/gene="NALP12"
/note="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP12"
/protein_id="AAO18163.1"
/db_xref="GI:28436378"

gene
CDS

ORIGIN

Query Match	96.5%;	Score	2997.8;	DB	9;	Length	3827;
Best Local Similarity	97.3%;	Pred. No. 0;					
Matches	3103;	Conservative	0;	Mismatches	2;	Indels	84;
						Gaps	2;
QY	1	ATGCTACGAACCGCAGCGAGGACGGCCCTCTGTCGGCTGTCCACCTTACTTGAAGAATCTC	60				
Db	177	ATGCTACGAACCGCAGCGAGGACGGCCCTCTGTCGGCTGTCCACCTTACTTGAAGAATCTC	236				
QY	61	GAGGCTGTGGAATCTGAAGAAGTTCAAGTTATATACCTGGGGACCGCGACAGAGCTGGGAGAA	120				
Db	237	GAGGCTGTGGAATCTGAAGAAGTTCAAGTTATATACCTGGGGACCGCGACAGAGCTGGGAGAA	296				
QY	121	GGCAAGATCCCTGGGGAGAGCATGGAAGAGCGCGGTCCTCGGAAATGCCCCAGCTGTCTC	180				
Db	297	GGCAAGATCCCTGGGGAGAGCATGGAAGAGCGCGGTCCTCGGAAATGCCCCAGCTGTCTC	356				
QY	181	ATCACCACTTCGGGCCAGAGGAGGCTCGAGGTTGGCTCTCAGACACTTTGAGCGGATA	240				
Db	357	ATCACCACTTCGGGCCAGAGGAGGCTCGAGGTTGGCTCTCAGACACTTTGAGCGGATA	416				
QY	241	AACAGGAAGACCTGTGGGAGAGGACAGACAGAGGACCTGGTGAGG	288				
Db	417	AACAGGAAGACCTGTGGGAGAGGACAGACAGAGAGGACCTGGTGAGGGATACCCACACT	476				
QY	289	-----	288				
Db	477	GGTGGCCGCTCTCACTTGGGAACCAAGTCAACATGCTCTTGGAACTCTCTCTGTCACT	536				
QY	289	-----GATCCCCAGAGAACTACAGGACTATGTCCGAGGAAATTCGGGCTCATG	339				
Db	537	CCAAGAAAGATCCCAGAGAACTACAGGACTATGTCCGAGGAAATTCGGGCTCATG	596				
QY	340	GAAGACCGCAATCGCGCTTAGGGGAATGTCAACCTCAGCCACCGGTACACCCGGCTC	399				
Db	597	GAAGACCGCAATCGCGCTTAGGGGAATGTCAACCTCAGCCACCGGTACACCCGGCTC	656				
QY	400	CTGCTGTGAAGAGACACTCAACCCCATGCAAGTTCAGAGCAGCTTCGGACAGAGC	459				
Db	657	CTGCTGTGAAGAGACACTCAACCCCATGCAAGTTCAGAGCAGCTTCGGACAGAGC	716				
QY	460	CGGGGACACGGAGGACCGTGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTT	519				
Db	717	CGGGGACACGGAGGACCGTGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTT	776				
QY	520	GAGCCAGCAGGAGCGCCCGAGCCACCGGCACCGTGTCTATGCAAGCGCGGAGGG	579				
Db	777	GAGCCAGCAGGAGCGCCCGAGCCACCGGCACCGTGTCTATGCAAGCGCGGAGGG	836				
QY	580	ATAGGCAAGTCCATGTGGCACAAGAGTGAATGCTGGACTGGGGGAGCGGAAGCTTTC	639				
Db	837	ATAGGCAAGTCCATGTGGCACAAGAGTGAATGCTGGACTGGGGGAGCGGAAGCTTTC	896				
QY	640	CAAGGCAGATTTGATTTATCTCTTCTACATCACTGAGGAGATGAACAGAGTGCCACG	699				
Db	897	CAAGGCAGATTTGATTTATCTCTTCTACATCACTGAGGAGATGAACAGAGTGCCACG	956				
QY	700	GAATGACAGCATGCAAGACCTCATCTTCAGCTGTGGCTTGAAGCCAGCGGCTCTCCAG	759				
Db	957	GAATGACAGCATGCAAGACCTCATCTTCAGCTGTGGCTTGAAGCCAGCGGCTCTCCAG	1016				
QY	760	GAGCTCATCGAGTTCGAGCGCTCTTTTATCATGACAGCGCTTCGATGAGCTCAAG	819				
Db	1017	GAGCTCATCGAGTTCGAGCGCTCTTTTATCATGACAGCGCTTCGATGAGCTCAAG	1076				
QY	820	CCTTTCTTCCAGATCTCAGGAGACCTGGTGCTCTGTCTGGAGAGAAACGGCCCAAG	879				
Db	1077	CCTTTCTTCCAGATCTCAGGAGACCTGGTGCTCTGTCTGGAGAGAAACGGCCCAAG	1136				
QY	880	GAGCTGTCTTTAAGCTTAATTCGGAAGAGCTGTCCCTGAGCTATCTTTGCTCATC	939				
Db	1137	GAGCTGTCTTTAAGCTTAATTCGGAAGAGCTGTCCCTGAGCTATCTTTGCTCATC	1196				
QY	940	ACCACAGGCCCCAGGCTTTGGAGAGCTCCACCGTCTGTGGAGCACCCAGGCATGTG	999				

1197	ACCACAGGCCACCGGCTTTGGAGAGCTCCACCGTCTGTGGAGCACCCCGAGCATGTG	1256
1000	GAGATCTCGGGCTTCTTGAGGCGAAGAGGAAGATTAATTCTTACAAGTATTTCCACAAT	1059
1257	GAGATCTCTGGGCTTCTCTGAGCGCAAGAGGAATATCTTCTAAGATATTTCCACAAT	1316
1060	GCAGAGCAGCGGGCCCAAGTCTTCAATTAACGTGAGGGACAACGAGCCCTCTTCCACCATG	1119
1317	GCAGAGCAGCGGGCCCAAGTCTTCAATTAACGTGAGGGACAACGAGCCCTCTTCCACCATG	1376
1120	TGCTTCGTCCCGCTGGTGTGTGGGTGTGTGTACTTCGCTCCAGCAGCAGCTGGAGGGT	1179
1377	TGCTTCGTCCCGCTGGTGTGTGGGTGTGTGTACTTCGCTCCAGCAGCAGCTGGAGGGT	1436
1180	GGGGGGCTGTTTGAGACAGACGTTCCAGGACCACTGCAAGTGTATCATGCTCTACTGCTG	1239
1437	GGGGGGCTGTTTGAGACAGACGTTCCAGGACCACTGCAAGTGTATCATGCTCTACTGCTG	1496
1240	AGTCTGATGCAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCCAGAGAGGG	1299
1497	AGTCTGATGCAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCCAGAGAGGG	1556
1300	TTGTGCTCTCTTGGCGGAGATGGGCTCTGGAAATCAGAAAATCCTATTTGAGCAGCAGGAC	1359
1557	TTGTGCTCTCTTGGCGGAGATGGGCTCTGGAAATCAGAAAATCCTATTTGAGCAGCAGGAC	1616
1360	CTCCGGAAAGCACGGCTTAGACGGGGAAGAGTCTCTGCTTCTCTCAACATGAACATCTTC	1419
1617	CTCCGGAAAGCACGGCTTAGACGGGGAAGAGTCTCTGCTTCTCTCAACATGAACATCTTC	1676
1420	CAGAGGACATCAACTGTGAGAGTACTACAGCTTCATCCACTTCAGTTTCCAGAAATTC	1479
1677	CAGAGGACATCAACTGTGAGAGTACTACAGCTTCATCCACTTCAGTTTCCAGAAATTC	1736
1480	TTTGTGAGCTATCTACTATCTCTGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1539
1737	TTTGTGAGCTATCTACTATCTCTGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1796
1540	GTGACGAGCTGTGTGACCGAGTACGCGTTTCTGAAAGAGAGTCTTCTGGCATCTCACGAG	1599
1797	GTGACGAGCTGTGTGACCGAGTACGCGTTTCTGAAAGAGAGTCTTCTGGCATCTCACGAG	1856
1600	CGTCTCTCTTTGGACTCTGAAACGAGGAGACAGGAGCCACTCGAGAGAGTCTCTCG	1659
1857	CGTCTCTCTTTGGACTCTGAAACGAGGAGACAGGAGCCACTCGAGAGAGTCTCTCG	1916
1660	TGGAAGCTCTCGCGCACATCAAGATGGACCTGTTGCACTGGATCCAAAGCAAGCTCAG	1719
1917	TGGAAGCTCTCGCGCACATCAAGATGGACCTGTTGCACTGGATCCAAAGCAAGCTCAG	1976
1720	AGCGACGGCTCCACCTCGAGCAGGGCTCTTGGAGTCTTTCAGCTGTGTACAGATC	1779
1977	AGCGACGGCTCCACCTCGAGCAGGGCTCTTGGAGTCTTTCAGCTGTGTACAGATC	2036
1780	CAGGAGGAGGTTTATCCAGAGCCCTGAGCCACTTCAGGTGATCTGTGTTCAGCAAC	1839
2037	CAGGAGGAGGTTTATCCAGAGCCCTGAGCCACTTCAGGTGATCTGTGTTCAGCAAC	2096
1840	ATTGCTCTCAAGATGAGACATGGTCTCTCTGTTCTGTCTGAAAGCGCTGCAGGAGCGCC	1899
2097	ATTGCTCTCAAGATGAGACATGGTCTCTCTGTTCTGTCTGAAAGCGCTGCAGGAGCGCC	2156
1900	CAGGTCTGCATTTGATGGCGCACTTACGCGCGGAACGGGGAAGAACCGCGGAGGTGC	1959
2157	CAGGTCTGCATTTGATGGCGCACTTACGCGCGGAACGGGGAAGAACCGCGGAGGTGC	2216
1960	TCCGAGGAGCGCACAGCTTGTGTGTCAGCTCAGACCCAGAGAGGACCGTTCTGTCTGGAC	2019
2217	TCCGAGGAGCGCACAGCTTGTGTGTCAGCT---ACCAGAGAGGACCGTTCTGTCTGGAC	2273
2020	GCCTACAGTGAACATCTGGCAGCGCGCTGTGCACCAATCCAAACCTGTATGAGCTGTCT	2079

```
Db 2274 GCCTACGTGAACATCTGGAGCGGCCCTGTGTGCCAACAATCCAAACCTGATAGAGCTGTCT 2333
QY 2080 CTGTACCGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGTCTGTCAAGGACTCAGACAC 2139
Db 2334 CTGTACCGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGTCTGTCAAGGACTCAGACAC 2393
QY 2140 CCAACTGCAAACTTTCAGAACTTGAGGCTGAAGAGTGCAGCTTCCTCCAGCTCAGCTGC 2199
Db 2394 CCAACTGCAAACTTTCAGAACTTGAGGCTGAAGAGTGCAGCTTCCTCCAGCTCAGCTGC 2453
QY 2200 GAGGACCTCTCTGAGCTCTCATAGCAATAGATTTCACAGAGTGCATCTCAGTGC 2259
Db 2454 GAGGACCTCTCTGAGCTCTCATAGCAATAGATTTCACAGAGTGCATCTCAGTGC 2513
QY 2260 AACGGCTGTGATTCACAGGATGATGTCTGTGGAGGGCTGGCGCATCCCCAGTGC 2319
Db 2514 AACGGCTGTGATTCACAGGATGATGTCTGTGGAGGGCTGGCGCATCCCCAGTGC 2573
QY 2320 AGGCTGAGATGATTCAGTTGAGGAGTGTGAGTGCAGTGGAGTCCGGGCTTGTGAGGAGATG 2379
Db 2574 AGGCTGAGATGATTCAGTTGAGGAGTGTGAGTGCAGTGGAGTCCGGGCTTGTGAGGAGATG 2633
QY 2380 GCTTCTGTGCTCGGACCAACCAATCTGTTGAGTGGAGTGCAGTGGAGTGCAGTGC 2439
Db 2634 GCTTCTGTGCTCGGACCAACCAATCTGTTGAGTGGAGTGCAGTGGAGTGCAGTGC 2693
QY 2440 GAGGATTTGGGCTGAGTTACTATGCCAGGAGTGCAGGACCCAGTCTGCAGATCAGG 2499
Db 2694 GAGGATTTGGGCTGAGTTACTATGCCAGGAGTGCAGGACCCAGTCTGCAGATCAGG 2753
QY 2500 ACTTGTGGCTGAAGATCTGCCGCTCTACTGTCTGTGCTGTGAGGAGTGCAGTGC 2559
Db 2754 ACTTGTGGCTGAAGATCTGCCGCTCTACTGTCTGTGCTGTGAGGAGTGCAGTGC 2813
QY 2560 CTCAGTGTGAACAGAGCTGAGAGCTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 2619
Db 2814 CTCAGTGTGAACAGAGCTGAGAGCTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 2873
QY 2620 GGGGTGTGCTGTGCTGTGAGGGCTCAGGATCCCACTGCAGTGCAGTGCAGTGCAGTGC 2679
Db 2874 GGGGTGTGCTGTGCTGTGAGGGCTCAGGATCCCACTGCAGTGCAGTGCAGTGCAGTGC 2933
QY 2680 TTGGCATCTGCCGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTGCTGTGCTGCAGGCC 2739
Db 2934 TTGGCATCTGCCGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTGCTGTGCTGCAGGCC 2993
QY 2740 AACCAACCTCCGGAGCTGAGTGTGAGTTCACGAGCTGGGAGCTGGGCTGTGG 2799
Db 2994 AACCAACCTCCGGAGCTGAGTGTGAGTTCACGAGCTGGGAGCTGGGCTGTGG 3053
QY 2800 TTGTGGCTGAGGGCTGCAACATCCCGCTGCAGACTCCAGAACTGTGCTGGATAGC 2859
Db 3054 TTGTGGCTGAGGGCTGCAACATCCCGCTGCAGACTCCAGAACTGTGCTGGATAGC 3113
QY 2860 TTGTGGCTGAGGGCTGAGAACTTTTCACTTCACTTGGGATCAACAGACC 2919
Db 3114 TTGTGGCTGAGGGCTGAGAACTTTTCACTTCACTTGGGATCAACAGACC 3173
QY 2920 TTGACCGACCTTTACCTGACCAACAGCCCTGAGGACACAGTGTCCGACTGCTTGC 2979
Db 3174 TTGACCGACCTTTACCTGACCAACAGCCCTGAGGACACAGTGTCCGACTGCTTGC 3233
QY 2980 AAGCGGTGAGCCATCTCGCTGCAAACTCCGAGTCTCTGTTATTTGGATGGAGCTG 3039
Db 3234 AAGCGGTGAGCCATCTCGCTGCAAACTCCGAGTCTCTGTTATTTGGATGGAGCTG 3293
QY 3040 AATAAATGACCCACAGTAGTGGAGCTTCGAGTAAACAAACCTTTATTTGGACATT 3099
Db 3294 AATAAATGACCCACAGTAGTGGAGCTTCGAGTAAACAAACCTTTATTTGGACATT 3353
QY 3100 GGCTGCTGA 3108
Db 3354 GGCTGCTGA 3362
```

```
RESULT 7
AY116205
LOCUS
DEFINITION
Homo sapiens monarch-1 splice form II mRNA, complete cds;
alternatively spliced.
ACCESSION
AY116205
VERSION
AY116205.1 GI:21711822
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3563)
AUTHORS
Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.
TITLE
Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat
Protein That Controls Classical and Nonclassical MHC Class I Genes
J. Immunol. 170 (11), 5354-5358 (2003)
JOURNAL
12/59408
PUBMED
2 (bases 1 to 3563)
REFERENCE
Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
AUTHORS
Direct Submission
TITLE
Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
JOURNAL
Road, Chapel Hill, NC 27599, USA
FEATURES
Location/Qualifiers
1..3563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
/cell_line="U937"
221..3241
/note="contains alternatively spliced exon 9"
CDS
/protein_id="AA05143.1"
/db_xref="GI:21711823"
/translation="MLRTAGDGLCLSTYLEEAEVLEAVELEKFKLYLGTATLGEKGIKIP
WGSMEKAGPLTEMAOLITHTFPEAWRLALSTFERINRKLWEGEDVLRDLPVPGG
PSSINGOSTCLLESLVTPRKDQPTVDYVRRKFRMLMEDNARLGEQDVLNLSHRTVL
LLVKEHNPVQOQQLDTCGHARTVGHQASPIKIEITLFEDEERPEPRITVQVGA
AGIKSMIAHKVMDLWDAGKLFQGRFDYLFYINCKEMQNSATECSMDLIFSCWPEPS
APQLRLRVPRLFIIDGDELKPSHPRHVEILGFSEAEKEYFYKYFHNAQAGQVFNVR
ELSLLITRPTALEKHLRLEHPRHVEILGFSEAEKEYFYKYFHNAQAGQVFNVR
DNEPLTFMCFVPLVWVCTCLOQLEGGLRQTSRTTAVYMLYLLSLMQPKGAP
RLOPPNORGLCSLAADGNOKILFEEDLRLKHLDGEDVSAFLNMIIFOKDINCER
YYSFIHLSFOEFFAAMYVILDEGGGAGDODVTRLLITYAFSERSFALISFLFGL
LNESTRHLEKSLCWKSPHIMDLQWISKAQSDGSTLOQSGLEFFSCLYEIQEE
FIQALSHFQIVVSNASKMEHVVSSFCRLKCRSAQVLDYATYSADGEDRACSA
GAHTLLVQLRPERTVLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCOGLRH
PNCKLQNLRLKRCRISSSACEDLSAALANKLTRMDLSGNGVFGPGMMLICEGLRHP
CRLQMLQRLKCOLESGACOMASVLTGTPHVLVDLTGNALEDLGLLCOGLRHPV
CRLTILWLKICRLTAACDELAETLSVNSQLRELDLSINELGDLGLVLLCSGLRHTC
KLQTLRIGICRLGSAACGLSVLQAHNLRELDSFNDLGDWGLWLLAEGLQHPACR
LQKLWLFGLDMLNDRTHSLAARVTKPYLDIGC"
869..892
/note="Region: P-loop"
2069..3238
/note="Region: leucine rich repeat"
221..289
/note="Region: pyrin domain"
```

ORIGIN

```
Query Match
Best Local Similarity 88.8%; Score 2759; DB 9; Length 3563;
Matches 2850; Conservative 0; Mismatches 0; Indels 81; Gaps 1;
Qy 1 ATGCTACGAACCGCAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
Db 221 ATGCTACGAACCGCAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 280
```

QY	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTTATACCTTGGGACCGGCACAGAGCTGGAGAA	120
DB	281	GAGGCTGTGAACTGAAGAAGTTCAAGTTTATACCTTGGGACCGGCACAGAGCTGGAGAA	340
QY	121	GGCAAGATCCCTTGGGGAAGCATGAGAAAGCCGCTCCCTTGGAAATGGCCACAGTGTCTC	180
DB	341	GGCAAGATCCCTTGGGGAAGCATGAGAAAGCCGCTCCCTTGGAAATGGCCACAGTGTCTC	400
QY	181	ATCACCCACTTGGGCGCAGAGAGGCTCGAGAGTTGGCTCTCAGCACCTTTTGGCGGATA	240
DB	401	ATCACCCACTTGGGCGCAGAGAGGCTCGAGAGTTGGCTCTCAGCACCTTTTGGCGGATA	460
QY	241	AACAGGAAGACCTGTGGAGAGAGACAGAGAGGACCTTGGTGAGG	288
DB	461	AACAGGAAGACCTGTGGAGAGAGACAGAGAGGACCTTGGTGAGGATACCCCACT	520
QY	289	-----	288
DB	521	GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCCCTTCTTGGAACTCTCTTTGTCACT	580
QY	289	-----GATCCCCAGAGAACTACAGGACTATGTCGCGAGGAAATTCGGGCTCATG	339
DB	581	CCAAGAAAAGATCCCCAGGAAACTTACAGGACTATGTCGCGAGGAAATTCGGGCTCATG	640
QY	340	GAAGACCGCAATCGCGCCCTAGGGGAATGTGCAACTCAGCCACCGGTACACCCGGTCTC	399
DB	641	GAAGACCGCAATCGCGCCCTAGGGGAATGTGCAACTCAGCCACCGGTACACCCGGTCTC	700
QY	400	CTGCTGTGTGAAGAGACACTCAAAACCCCATGAGTCCAGCAGCAGCTTCTGACACAGGC	459
DB	701	CTGCTGTGTGAAGAGACACTCAAAACCCCATGAGTCCAGCAGCAGCTTCTGACACAGGC	760
QY	460	CGGGACACACGAGGACCGTGGGACACCAAGCTAGCCCATCAAGATAGAGACCTCTTT	519
DB	761	CGGGACACACGAGGACCGTGGGACACCAAGCTAGCCCATCAAGATAGAGACCTCTTT	820
QY	520	GAGCCAGACGAGGACGCCCCCGACCCGCGCACCTGGTTCATGCAAGCGCGGACAGG	579
DB	821	GAGCCAGACGAGGACGCCCCCGACCCGCGCACCTGGTTCATGCAAGCGCGGACAGG	880
QY	580	ATAGGCAAGTCCATGTGTGGACACCAAGGTGATCTGGACTGGGCGGACGGGAAGTCTTTC	639
DB	881	ATAGGCAAGTCCATGTGTGGACACCAAGGTGATCTGGACTGGGCGGACGGGAAGTCTTTC	940
QY	640	CAAGGCAGATTGATATCTCTTCTACATCAACTGCGAGGAGATGAACAGAGTGCACG	699
DB	941	CAAGGCAGATTGATATCTCTTCTACATCAACTGCGAGGAGATGAACAGAGTGCACG	1000
QY	700	GAATGCAGCATGCAAGACCTCATCTTCAGCTGTGGCCTGAGCCGACGCGCTCTCCAG	759
DB	1001	GAATGCAGCATGCAAGACCTCATCTTCAGCTGTGGCCTGAGCCGACGCGCTCTCCAG	1060
QY	760	GAGCTCATCCGAGTTCGAGGCGCTCCTTTTCAATCAATCGACGGCTTGGATGAGTCAAG	819
DB	1061	GAGCTCATCCGAGTTCGAGGCGCTCCTTTTCAATCAATCGACGGCTTGGATGAGTCAAG	1120
QY	820	CTTTCTTTCCAGATCTCTCAGGACCTGCTGCTGCTGGGAGGAGAAACGGCCACG	879
DB	1121	CTTTCTTTCCAGATCTCTCAGGACCTGCTGCTGCTGGGAGGAGAAACGGCCACG	1180
QY	880	GAGCTGCTTTTAAACAGTTAATTCGGAAGAAGTGTCTCCCTGAGCTATCTTTTGTCTATC	939
DB	1181	GAGCTGCTTTTAAACAGTTAATTCGGAAGAAGTGTCTCCCTGAGCTATCTTTTGTCTATC	1240
QY	940	ACCACAGGCCCCAGGCTTTGGAGAAGCTCCACGCTGTGTGGAGCACCCGAGCATGTG	999
DB	1241	ACCACAGGCCCCAGGCTTTGGAGAAGCTCCACGCTGTGTGGAGCACCCGAGCATGTG	1300
QY	1000	GAGATCTTGGCTTCTCTGAGGACGAAGGAGAAATCTTCTACAGTATTTCCCAAT	1059
DB	1301	GAGATCTTGGCTTCTCTGAGGACGAAGGAGAAATCTTCTACAGTATTTCCCAAT	1360
QY	1060	GCAGAGCAGGCGGGCCAAAGTCTTCAATTAAGTGAAGGACCAAGAGCTCTCTTCCACATG	1119

Db	1361	GCAGACGCGCGGCGCAAGTCTTCAATTACGTGAGGACAACGAGCCTCTCTTACCATG	1420
Qy	1120	TGCTTCGTCCCCCTGGTGTGCTGGGTGTGTACTCGCTCCAGCAGCAGCTGGAGGT	1179
Db	1421	TGCTTCGTCCCCCTGGTGTGCTGGGTGTGTACTCGCTCCAGCAGCAGCTGGAGGT	1480
Qy	1180	GGGGGCTGTGTAGACAGACGCTCCAGGACCACTACGACGTGATCATGCTCTACTCGTGTG	1239
Db	1481	GGGGGCTGTGTAGACAGACGCTCCAGGACCACTACGACGTGATCATGCTCTACTCGTGTG	1540
Qy	1240	AGTCTGATCGAACCCAAAGCCGGGGGCGCGCGCTCCAGCCCCCAACCAACAGAGAGGG	1299
Db	1541	AGTCTGATCGAACCCAAAGCCGGGGGCGCGCGCTCCAGCCCCCAACCAACAGAGAGGG	1600
Qy	1300	TTTGCTCTCTTTGGCGCAGATGGGCTCTGGAATCAGAAAAATCCTATTTGAGGACGAGAC	1359
Db	1601	TTTGCTCTCTTTGGCGCAGATGGGCTCTGGAATCAGAAAAATCCTATTTGAGGACGAGAC	1660
Qy	1360	CTCCGAAGACGCGCTAGACGGGGGAAGCGTCTCTGCGCTTCTCTCAATGAAATCTTCTC	1411
Db	1661	CTCCGAAGACGCGCTAGACGGGGGAAGCGTCTCTGCGCTTCTCTCAATGAAATCTTCTC	1720
Qy	1420	CAGAAAGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATTC	1479
Db	1721	CAGAAAGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATTC	1780
Qy	1480	TTTGCACTATGTACTATATCTCTGACGAGGGGGGCGGGCCAGACCAAGGAC	1539
Db	1781	TTTGCACTATGTACTATATCTCTGACGAGGGGGGCGGGCCAGACCAAGGAC	1840
Qy	1540	GTGACGAGCTGTGACCGGATACGGTTTTCTGAAAGAGGCTTCTGTGCATCCACCAG	1599
Db	1841	GTGACGAGCTGTGACCGGATACGGTTTTCTGAAAGAGGCTTCTGTGCATCCACCAG	1900
Qy	1600	CGCTTCCTGTTTGGACCTCTGACGAGGAGACAGGAGCCACTCGAGAAAGAGTCTCTGC	1659
Db	1901	CGCTTCCTGTTTGGACCTCTGACGAGGAGACAGGAGCCACTCGAGAAAGAGTCTCTGC	1960
Qy	1660	TGGAAGTCTCGCGGCACATCAAGATGGACCTCTTGCACTGGATCCAAAGCAAGCTCAG	1719
Db	1961	TGGAAGTCTCGCGGCACATCAAGATGGACCTCTTGCACTGGATCCAAAGCAAGCTCAG	2020
Qy	1720	AGCGACGGCTCCACCTGACGAGGGCTCTTGGAGTTCTTCACTGCTGTTACGAGATC	1779
Db	2021	AGCGACGGCTCCACCTGACGAGGGCTCTTGGAGTTCTTCACTGCTGTTACGAGATC	2080
Qy	1780	CAGGAGAGGAGTTATCCAGCAGGCGCTGAGCCACTTCCAGTGTATCGTGTGAGCAAC	1839
Db	2081	CAGGAGAGGAGTTATCCAGCAGGCGCTGAGCCACTTCCAGTGTATCGTGTGAGCAAC	2140
Qy	1840	ATTGCTCCAAAGATGGAGCATGGTCTCCTCGTCTTGTGTAAGGGCTGCAGGAGCGCC	1899
Db	2141	ATTGCTCCAAAGATGGAGCATGGTCTCCTCGTCTTGTGTAAGGGCTGCAGGAGCGCC	2200
Qy	1900	CAGGTGCTGCACTTGTATGGCGCCACTACAGCGCGGACGGGGAAGACCGCGGAGGTGC	1959
Db	2201	CAGGTGCTGCACTTGTATGGCGCCACTACAGCGCGGACGGGGAAGACCGCGGAGGTGC	2260
Qy	1960	TCCGAGGAGGCGACACGCTGTTGGTGCAGCTCAGACCAGAGAGGACCGTTCCTGCTGGAC	2019
Db	2261	TCCGAGGAGGCGACACGCTGTTGGTGCAGCTCAGACCAGAGAGGACCGTTCCTGCTGGAC	2320
Qy	2020	GCCTACAGTGAACATCTGGCAGCGGCGCTGTGACCAATCCAACTGTATAGAGCTGTCT	2079
Db	2321	GCCTACAGTGAACATCTGGCAGCGGCGCTGTGACCAATCCAACTGTATAGAGCTGTCT	2380
Qy	2080	CTGTACCGAAATGCCCTGGGACGCCGGGGGTGAAGTGTCTGTCAAGGACTCAGACAC	2139
Db	2381	CTGTACCGAAATGCCCTGGGACGCCGGGGGTGAAGTGTCTGTCAAGGACTCAGACAC	2440
Qy	2140	CCCAACTGCAAACTTCAGAACCTTAGGCTGAAGGTTGCCGATCTCCAGCTCAGCCTGC	2199

[illegible]

Db 1465 AGTCTGATGCAACCAAGCCGGGGGCCCCCGGCTTCAGCCGCCCAACCAAGAGAGG 1524
Qy 1300 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1359
Db 1525 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1584
Qy 1360 CTCGGAAGCAGCGCTAGACGGGGAAGACGTCCTGCTGCTTCTCTCAACATGAAATCTTC 1419
Db 1585 CTCGGAAGCAGCGCTAGACGGGGAAGACGTCCTGCTTCTCTCAACATGAAATCTTC 1644
Qy 1420 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTTCATCCATTTGAGTTTCCAGGAATTC 1479
Db 1645 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTTCATCCATTTGAGTTTCCAGGAATTC 1704
Qy 1480 TTTCAGCTATGTAATATATCTTGAGAGGGGAGGCGGGGAGGCGCCAGACAGGAC 1539
Db 1705 TTTCAGCTATGTAATATATCTTGAGAGGGGAGGCGGGGAGGCGCCAGACAGGAC 1764
Qy 1540 GTGACCAAGGCTTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTGGCACTCACAGC 1599
Db 1765 GTGACCAAGGCTTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTGGCACTCACAGC 1824
Qy 1600 CGCTTCTGTTTGGACTCTGAAACGAGGAGACAGGAGCCACTGGAGAGAGTCTCTGC 1659
Db 1825 CGCTTCTGTTTGGACTCTGAAACGAGGAGACAGGAGCCACTGGAGAGAGTCTCTGC 1884
Qy 1660 TGAAGGTCTGCGCGCATCAAGATGGACCTGTTGCACTGGATCCAAAGCAAGCTCAG 1719
Db 1885 TGAAGGTCTGCGCGCATCAAGATGGACCTGTTGCACTGGATCCAAAGCAAGCTCAG 1944
Qy 1720 AGCGAGGCTCCACCGCTGACAGGGCTCTGAGGTTCTTCACTGCTTCTGAGATC 1779
Db 1945 AGCGAGGCTCCACCGCTGACAGGGCTCTGAGGTTCTTCACTGCTTCTGAGATC 2004
Qy 1780 CAGGAGGAGGAGTTTATCCAGAGCGCTGAGGACATTCAGAGTGATCGTGGTCAAGAAC 1839
Db 2005 CAGGAGGAGGAGTTTATCCAGAGCGCTGAGGACATTCAGAGTGATCGTGGTCAAGAAC 2064
Qy 1840 ATTGCTTCAAGATGAGAGCAATGGTCTCTGCTGTTCTGTTGAAGCGCTGAGAGGCGCC 1899
Db 2065 ATTGCTTCAAGATGAGAGCAATGGTCTCTGCTGTTCTGTTGAAGCGCTGAGAGGCGCC 2124
Qy 1900 CAGTGTGCACTTGTATGGCGCACCTTACAGCGGAGCGGAGAGCGCGCGAGGTGC 1959
Db 2125 CAGTGTGCACTTGTATGGCGCACCTTACAGCGGAGCGGAGAGCGCGCGAGGTGC 2184
Qy 1960 TCCGAGAGGCGCACAGCTGTTGTGAGCTCAGACACAGAGAGACCGTTCTGTGGAC 2019
Db 2185 TCCGAGAGGCGCACAGCTGTTGTGAGCT--ACAGAGAGGACCGTTCTGTGGAC 2241
Qy 2020 GCCTACAGTGAACATCTGGCAGCGGCTGTGACCAATCCAAACCTGATAGAGTGTCT 2079
Db 2242 GCCTACAGTGAACATCTGGCAGCGGCTGTGACCAATCCAAACCTGATAGAGTGTCT 2301
Qy 2080 CTGTACCGAATCCCTGGGAGCGGGGCTGAGCTGCTGCAAGGACTCAGACAC 2139
Db 2302 CTGTACCGAATCCCTGGGAGCGGGGCTGAGCTGCTGCAAGGACTCAGACAC 2361
Qy 2140 CCCAACTGCAAACTTCAGAACTCAGGCTGAAGAGGTCGCGCATCTCCAGCTCAGCGTGC 2199
Db 2362 CCCAACTGCAAACTTCAGAACTCAGGCTGAAGAGGTCGCGCATCTCCAGCTCAGCGTGC 2421
Qy 2200 GAGGACCTCTCTGAGCTCTCATGACCAATTAAGAAATTTGAAGATGGATCTCAGTGGC 2259
Db 2422 GAGGACCTCTCTGAGCTCTCATGACCAATTAAGAAATTTGAAGATGGATCTCAGTGGC 2481
Qy 2260 AACGGGCTGATTTCCAGGATGATGCTGCTTTGCGAGGGCTCGCGCATCCCGAGTGC 2319
Db 2482 AACGGGCTGATTTCCAGGATGATGCTGCTTTGCGAGGGCTCGCGCATCCCGAGTGC 2541
Qy 2320 AGGCTGAGATGATTTCAAGTTGAGGAAGTGTGAGCTGAGTCCGGGGCTTGTGAGGATG 2379
Db 2542 AGGCTGAGATGATTTCAAGTTGAGGAAGTGTGAGCTGAGTCCGGGGCTTGTGAGGATG 2601

Qy 2380 GCTTCTGTCTGCGACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2439
Db 2602 GCTTCTGTCTGCGACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2661
Qy 2440 GAGGATTTGGGCTGAGGTTACTATGACAGGACTGAGGACCCAGTCTGAGACTACGG 2499
Db 2662 GAGGATTTGGGCTGAGGTTACTATGACAGGACTGAGGACCCAGTCTGAGACTACGG 2721
Qy 2500 ACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGTGACGAGCTGGCCCTCAACT 2559
Db 2722 ACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGTGACGAGCTGGCCCTCAACT 2781
Qy 2560 CTCAGTGTGAACAGAGGCTGAGAGAGCTGAGACTGAGCTGAATGAGCTGGGGACCTC 2619
Db 2782 CTCAGTGTGAACAGAGGCTGAGAGAGCTGAGACTGAGCTGAATGAGCTGGGGACCTC 2841
Qy 2620 GGGGTGCTGCTGCTGTGAGGGCTCAGGACATCCAGCTGCAAGCTCCAGACCTTCGG 2679
Db 2842 GGGGTGCTGCTGCTGTGAGGGCTCAGGACATCCAGCTGCAAGCTCCAGACCTTCGG 2899
Qy 2680 TTGGGCACTCTCCGCGCTGGGCTCTGCGGCTGTGAGGGTCTTCTGTGCTGCTCAGGCC 2739
Db 2900 ----- 2899
Qy 2740 AACCAACCTCCGAGAGCTGGAATTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
Db 2900 ----- 2899
Qy 2800 TTGCTGCTGAGGGCTGCAACATCCCGCTCAGACTCCAGAACTGTGGCTGGATAGC 2859
Db 2900 ----- 2910
Qy 2860 TGTGGCTCAGACCAAGGCTTGTGAGAATCTTACTTACCTGGGATCAACAGACCTC 2919
Db 2911 TGTGGCTCAGACCAAGGCTTGTGAGAATCTTACTTACCTGGGATCAACAGACCTC 2970
Qy 2920 TTGACGAGCTTTACTGACCAACAAAGCCCTAGGGAGACAGGTGTCCGACTGCTTTC 2979
Db 2971 TTGACGAGCTTTACTGACCAACAAAGCCCTAGGGAGACAGGTGTCCGACTGCTTTC 3030
Qy 2980 AAGCGCTGAGCCATCTGCTGCAAACTCGAGTCTCTGCTTATTTGGGATGACCTG 3039
Db 3031 AAGCGCTGAGCCATCTGCTGCAAACTCGAGTCTCTGCTTATTTGGGATGACCTG 3090
Qy 3040 AATAAATGACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAAACCTTATTTGACATT 3099
Db 3091 AATAAATGACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAAACCTTATTTGACATT 3150
Qy 3100 GGCTGCTGA 3108
Db 3151 GGCTGCTGA 3159

RESULT 10
AY116206

LOCUS Homo sapiens monarch-1 splice form III mRNA, complete cds;
DEFINITION AY116206 3395 bp mRNA linear PRI 21-MAY-2003
alternatively spliced.

ACCESSION AY116206
VERSION AY116206.1 GI:21711824

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3395)

AUTHORS Williams K.L., Taxman D.J., Linhoff M.W., Reed W. and Ting J.P.Y.

TITLE Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes

JOURNAL J. Immunol. 170 (11), 5354-5358 (2003)

PUBMED 12759408

REFERENCE 2 (bases 1 to 3395)

AUTHORS

Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.

TITLE

Direct Submission
Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA

FEATURES

Location/Qualifiers

source

1. 3395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221..3073
/notes="contains alternatively spliced exons 7 and 8"

CDS

/codon_start=1
/product="monarch-1 splice form III"
/protein_id="AAW75144.1"
/db_xref="GI:21711825"
/translation="MURTAGRDLRLSTVBLEAVELEKFKLYLGTATELGEKIP
WSEKAGLEMAQLLITHTFPEEAWRLALSTFERINRDLWREGQEDLVRDTPPG
PSLGNQSTCLLEVLTPKQDPOETRYDVRKFRIMEDRNARLCEVNLSHRYRL
LLVKEHNPQVQOQLDTRGHARTVGHQASPIKETLFEDEERPEPRTVMQGA
AGIGKMLAHKVLMDKGLFOGRDYLFYINCREMNSATECSMDLIFSCWPEPS
APLOELLRVERLLFIIDFDELKPSFDPQGGWCLCWEKRPTELLLSLRKLLP
EUSLLITPTALEKHLRLEHPRHVEILGFEAEKKEYFYKHNAEQGVFNPYR
DNEPLFTMCVPLVCCVCTQOQLEGGGLRQTSRTTAVYMLYLLSLMOPKPGAP
RYOPPQRGLSLAADGLWNOKILFEEODLRKHLDGDSVAFSLNMNI FOKDINER
YYSFIHLSQEPFLAAMYIYLDDEGEGAGPDQDVTLLTEYAPERSFLALTRFLPL
LNEERSHLEKSLCWKVSPIKMDLLQWTSKAQSGSTLQOGLSLEFFSLVEIQREE
FTQALLSHQVIVSNIAKSEHWSFCLKRCRQAQVHLHYATYSADGEDRACRS
GAQLLVQRPERTVLIDYSEHLAALCTNPNI ELSYRNALSGRKYLLCQGLRH
PCKLQNRKRCRISSEACEDLSAALIANKNLTRMDLSCNGVGFPMMLLCEGLRP
QCLQMIQURKQLESGAQEMASVLGTNPHLVELDLTGNALEDLGLRLCCGLRHPV
CKRLTLMWLDSCGLTAKACENLYFTLGINOTLTDLYLNNALDGTGVRLCKRLSHP
869..892
/note="Region: P-loop"
2069..3070
/note="Region: leucine rich repeat"
221..289
/note="Region: pyrin domain"

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 79.5%; Score 2470; DB 9; Length 3395;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 65; Indels 87; Gaps 2;

QY 1 ATGCTACGAAACCGCAGGAGGAGCGCCCTCTGCGCTGTCCACCTACTTGGAGAACTC 60
DB 221 ATGCTACGAAACCGCAGGAGGAGCGCCCTCTGCGCTGTCCACCTACTTGGAGAACTC 280

QY 61 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCAGACAGAGCTGGGAGAA 120
DB 281 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCAGACAGAGCTGGGAGAA 340

QY 121 GCAAGATCCCTGGGAGAGCATGGAAGAGCGGCTCCCTCGAATGGCCAGCTGCTC 180
DB 341 GCAAGATCCCTGGGAGAGCATGGAAGAGCGGCTCCCTCGAATGGCCAGCTGCTC 400

QY 181 ATCAACCATTCGGCCAGAGAGGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
DB 401 ATCAACCATTCGGCCAGAGAGGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 460

QY 241 AACAGAGAGCATGTGGGAGAGAGCAGAGAGGACCTGTGAGG----- 288
DB 461 AACAGAGAGCATGTGGGAGAGAGCAGAGAGGACCTGTGAGGATACCCACCT 520

QY 289 ----- 288
DB 521 GGTGGCCGCTCCTCACTTGGGAACAGTCAACATGCTTCTTGAAGTCTCTTGTCACT 580

QY 289 -----GATCCCAGGAAACCTACAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 339
DB 581 CCAAGAAAAGATCCCAGGAAACCTACAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 640

QY 1420 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATTC 1479
Db 1721 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATTC 1780
QY 1480 TTTTGAGCTATGTACTATATCTGACAGAGGGGAGGGCGGGAGGCCAGGACAGGAC 1539
Db 1781 TTTTGAGCTATGTACTATATCTGACAGAGGGGAGGGCGGGAGGCCAGGACAGGAC 1840
QY 1540 GTGACAGGCTGTGACAGGATGACGCTTTTCTGAAAGAGGCTTCTTGCACTCAACGAC 1599
Db 1841 GTGACAGGCTGTGACAGGATGACGCTTTTCTGAAAGAGGCTTCTTGCACTCAACGAC 1900
QY 1600 CGCTTCCTGTTGAGTCTCTGACAGGAGGACAGGACAGGACAGGACAGGACAGGAC 1659
Db 1901 CGCTTCCTGTTGAGTCTCTGACAGGAGGACAGGACAGGACAGGACAGGACAGGAC 1960
QY 1660 TGAAGGTCTCGCCGACATCAAGATGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 1719
Db 1961 TGAAGGTCTCGCCGACATCAAGATGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 2020
QY 1720 AGCGAGGCTCCACCTCGAGAGGCTCTCTGAGGTTCTTGAAGTCTTGTGACGATC 1779
Db 2021 AGCGAGGCTCCACCTCGAGAGGCTCTCTGAGGTTCTTGAAGTCTTGTGACGATC 2080
QY 1780 CAGGAGGAGGATTTATCCAGAGGCTCTGAGGACCTTCCAGGTCATGCTGTCAGCAC 1839
Db 2081 CAGGAGGAGGATTTATCCAGAGGCTCTGAGGACCTTCCAGGTCATGCTGTCAGCAC 2140
QY 1840 ATTGCTCTCAAGATGAGGACATGCTCTGCTGTTGAGGTTCTTGAAGGCTCGAGAGCGCC 1899
Db 2141 ATTGCTCTCAAGATGAGGACATGCTCTGCTGTTGAGGTTCTTGAAGGCTCGAGAGCGCC 2200
QY 1900 CAGGTGCTGACATTTGATGGCCACCTTACAGCGCGGACGGGAGAGACCGCGCGAGGTGC 1959
Db 2201 CAGGTGCTGACATTTGATGGCCACCTTACAGCGCGGACGGGAGAGACCGCGCGAGGTGC 2260
QY 1960 TCCGAGGAGGACACGCTGTTGTCAGCTCAGACGAGGACGACCGTCTGCTGAC 2019
Db 2261 TCCGAGGAGGACACGCTGTTGTCAGCTCAGACGAGGACGACCGTCTGCTGCTGAC 2320
QY 2020 GCCTACAGTGAACATCTGCGACGCGCCCTGTGACCAATCCAACTGATAGAGTGTCT 2079
Db 2321 GCCTACAGTGAACATCTGCGACGCGCCCTGTGACCAATCCAACTGATAGAGTGTCT 2380
QY 2080 CTGTACCGAAATGCTCGGACGCGCGGGGTGAAGTGTCTGTGCAAGGACTCAGACAC 2139
Db 2381 CTGTACCGAAATGCTCGGACGCGCGGGGTGAAGTGTCTGTGCAAGGACTCAGACAC 2440
QY 2140 CCCAAGTCAAACTTCAGAACTGAGGCTGAAGGTGCGCATCTCCAGCTCAGCCTGC 2199
Db 2441 CCCAAGTCAAACTTCAGAACTGAGGCTGAAGGTGCGCATCTCCAGCTCAGCCTGC 2500
QY 2200 GAGGACCTCTCTGAGCTCTCATAGCCAAATGAAGATTTGACAAAGATGATCTCAGTGGC 2259
Db 2501 GAGGACCTCTCTGAGCTCTCATAGCCAAATGAAGATTTGACAAAGATGATCTCAGTGGC 2560
QY 2260 AACGGCTGTGATTTCCAGGATGATGCTGTTGTCGAGGCTGCGGATCCCGAGTGC 2319
Db 2561 AACGGCTGTGATTTCCAGGATGATGCTGTTGTCGAGGCTGCGGATCCCGAGTGC 2620
QY 2320 AGGCTGAGATGATTTGAGTGAAGGATGTCAGCTGAGTCCGGGCTTGTGAGGATG 2379
Db 2621 AGGCTGAGATGATTTGAGTGAAGGATGTCAGCTGAGTCCGGGCTTGTGAGGATG 2680
QY 2380 GCTTCTGTGCTCGGACCAACCACTGTTGATTTGAGTGGACCTGACAGGAAATGCACTG 2439
Db 2681 GCTTCTGTGCTCGGACCAACCACTGTTGATTTGAGTGGACCTGACAGGAAATGCACTG 2740
QY 2440 GAGGATTTGGGCTCAGGTTACTATGACAGGACCTGAGGACCCAGTCTGACACTACGG 2499
Db 2741 GAGGATTTGGGCTCAGGTTACTATGACAGGACCTGAGGACCCAGTCTGACACTACGG 2800
QY 2500 ACTTGTGGCTGAAGAT-----CTGCGGCTCTGCTGCTGCTGTCAGGAGTGGCC 2553

Db 2801 ACTTGTGGCTGTGGCTGATAGCTGTGGCTTCACAGCAAGGCTTGTGAGATTTTAC 2860
QY 2554 TCAACTCTCAGTGTGAACAGGAGCTGAGAGCTGAGACCTGAGCTGAATGAGTGGGG 2613
Db 2861 TTAACCTTGGGATCAACAGGACCTTGAACGACCTTACCTGACCAACAGCCCTAGGG 2920
QY 2614 GACCTCGGGGTGCTGCTGCTGAGGCTTCAAGCATCCACCTGCAAGCTCCAGACC 2673
Db 2921 GACACAGGTGCTGCTGCTTGAAGGCTGAGCCTTCTGGCTGCAAACTCCGAGTC 2980
QY 2674 CTGCGGTT 2681
Db 2981 CTCTGTT 2988

RESULT 11
AY116207
LOCUS AY116207 Homo sapiens monarch-1 splice form IV mRNA, complete cds;
DEFINITION alternatively spliced.
ACCESSION AY116207
VERSION AY116207.1 GI:21711826
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3221)
AUTHORS Williams K.L., Taxman, D.J., Linhoff M.W., Reed W. and Ting, J.P.Y.
TITLE Monarch-1: A Porphyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes
J. Immunol. 170 (11), 5354-5358 (2003)
PUBMED 12759408
REFERENCE 2 (bases 1 to 3221)
AUTHORS Williams K.L., Linhoff, M.W. and Ting, J.P.Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA
FEATURES
1. 3221
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221..2899
/note="contains alternatively spliced exons 7, 8 and 9"
CDS
/codon_start=1
/product="monarch-1 splice form IV"
/protein_id="AA075145.1"
/db_xref="GI:21711827"
/translation="MLRTAGRDGLCRSLTYLEAEVLEKFKFLVLTATBLGKIP
WGSMEKAGPLEMAOLLTHFGPEAWRLALSTFRINPKDIEWGOREDIARDTPPGG
PSSVNGSTCLVSLVTPKDPQTYRDYVRKFRLEMDNABLCGVNLSHRYTRL
LLVKEHNSPMQVQQLDTRGHARTVGHQSPKIKETLFPDEERPEPPRVNMQGA
AGIKSLMAHKLMDWADKLFQGRFDYLFYINCRMNQATECSMODLIFSCWPEPS
APQLLRVPERLLIFIDGDELKPSFDPQWCLWBEKRPTELNLNLSIRKLLP
ELSLITRPTALEKLHRLHPRHVEILGFSEAEKEYFYKFNABQAGQVFNVR
DNEPLTFMCPVLVWVCTCLOQLEGGLRDTSTTAVVMLYLLSLMQPKGAP
RLOPPNORGLCSLAADGLWOKILFPEODLRKHGLDGEDYSALNNMIFOKDINCR
YYSIHLSPQEFPAAMYIILDEGGAGPQDDVTRLLTEYAFSEFSLIISRLFGL
LNESTRHLEKSLCWKSPHIMDLOQWISQDSTLQOGLSFELFSLYFIOEE
FIQALSHFQIVVSNIAKMEHVSFCLKRCRQAQVHLHYGATYAGSDGRARCS
FAHTLVQLRPERVTLDAVSEHLAALCTNPILSLYRNALSGVGLKLCQGLRH
PNCQLNLRKCRISACDLSAALIANKLRMDLSNGVGPFGMMLCEGLRHP
CQRLQMLKQLESACQEMASVLGTNPHLVELDLTGNALEDLGLRLICQGLRHP
CRLLTLWLFQMDLNKMTSHSLAALRVTKPILDIGC"
869..892
/note="Region: P-loop"
2069..2896
/note="Region: leucine rich repeat"
221..2899
misc_feature
869..892
misc_feature
2069..2896
misc_feature
221..2899

```
/note="Region: pyrin domain"
```

Query Match 77.8%; Score 2417; DB 9; Length 3221;
Best Local Similarity 96.9%;
Matches 2508; Conservative 0; Mismatches 0; Indels 81;
Pred. No. 0;

QY	1	ATGCTACGAACCGGACGAGGACGGCCCTCTGTGCGCTGTCCAACTACTTGAAGAACTC	60
Db	221	ATGCTACGAACCGGACGAGGACGGCCCTCTGTGCGCTGTCCAACTACTTGAAGAACTC	280
QY	61	GAGCGTGTGGAACCTGAAGAGTTTCAAGTTATACCTTGGGACCGCGACACAGCTGGGAA	120
Db	281	GAGCGTGTGGAACCTGAAGAGTTTCAAGTTATACCTTGGGACCGCGACACAGCTGGGAA	340
QY	121	GGCAAGATCCCTCTGGGGAAGCATGGAGAAAGCCGCTCCCTTGGAAATGGCCACAGCTGCTC	180
Db	341	GGCAAGATCCCTCTGGGGAAGCATGGAGAAAGCCGCTCCCTTGGAAATGGCCACAGCTGCTC	400
QY	181	ATACCCCACTTCGGGCCAGAGGAGCCCTGGAGTTGGCTCTCAGCACTTTGAGCGGATA	240
Db	401	ATACCCCACTTCGGGCCAGAGGAGCCCTGGAGTTGGCTCTCAGCACTTTGAGCGGATA	460
QY	241	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTTGGTGAGG-----	288
Db	461	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTTGGTGAGGATACCCCACCT	520
QY	289	-----	288
Db	521	GGTGGCCCGCTCCTCACTTGGGAACCACTCAACATGCGCTTCTGGAAGTCTCTCTTGTCACT	580
QY	289	-----GATCCCCAGGAACCTACAGGAGACTATGTCCGACGAAATCCCGGCTCATG	339
Db	581	CCAAGAAAAGATCCCCAGGAACCTACAGGAGACTATGTCCGACGAAATCCCGGCTCATG	640
QY	340	GAAGACCGCAATGGCGCCTTAGGGAAATGTGTAACTCAGCCACCCGGTACACCCGGCTC	399
Db	641	GAAGACCGCAATGGCGCCTTAGGGAAATGTGTAACTCAGCCACCCGGTACACCCGGCTC	700
QY	400	CTGCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGAGCAGCAGCTTCTGGACACAGGC	459
Db	701	CTGCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGAGCAGCAGCTTCTGGACACAGGC	760
QY	460	CGGGGACACGGGAGGACCGTGGGACACCAAGCTAGCCCTATCAAGATAGAGAACCCTCTTT	519
Db	761	CGGGGACACGGGAGGACCGTGGGACACCAAGCTAGCCCTATCAAGATAGAGAACCCTCTTT	820
QY	520	GAGCCACAGAGGAGCGCCCCGAGCCACCGCGACCGTGTCTATGCAAGCGCGGCGAGGG	579
Db	821	GAGCCACAGAGGAGCGCCCCGAGCCACCGCGACCGTGTCTATGCAAGCGCGGCGAGGG	880
QY	580	ATAGGCAAGTCCATGTGTGGACACAAGGTTGATGTGACTGGGCGGACGGGAAGCTCTTC	639
Db	881	ATAGGCAAGTCCATGTGTGGACACAAGGTTGATGTGACTGGGCGGACGGGAAGCTCTTC	940
QY	640	CAAGGCAGATTTCATTATCTCTTCTACATCAACTGAGGAGGATGACACAGATGCGCACG	699
Db	941	CAAGGCAGATTTCATTATCTCTTCTACATCAACTGAGGAGGATGACACAGATGCGCACG	1000
QY	700	GAATGACGATGTCAGGAGCGCTCTTTTCATCATCGACGGTTCGATGAGCTCAAG	759
Db	1001	GAATGACGATGTCAGGAGCGCTCTTTTCATCATCGACGGTTCGATGAGCTCAAG	1060
QY	760	GAGCTCATCGAGTTCCTCGAGCGCTCTTTTCATCATCGACGGTTCGATGAGCTCAAG	819
Db	1061	GAGCTCATCGAGTTCCTCGAGCGCTCTTTTCATCATCGACGGTTCGATGAGCTCAAG	1120
QY	820	CTTCTTTTCCACGATCCTCAGGAGCCCTGGTGCTCTGTGGGAGGAGAAACGGCCACG	879
Db	1121	CTTCTTTTCCACGATCCTCAGGAGCCCTGGTGCTCTGTGGGAGGAGAAACGGCCACG	1180
QY	880	GAGCTGCTTCTTAAACAGCTTAATTGGGAAGACGCTGCTCCCTGAGCTATCTTTGCTCATC	939

QY	889	CTTTAA	CAGCTTAATTCGGAA	GAAGCTGCTCCCTG	SAGCTATCTTTGCTCATCACCA	CGG	948
DB	79421	CTTTAA	CAGCTTAATTCGGAA	GAAGCTGCTCCCTG	SAGCTATCTTTGCTCATCACCA	CGG	79362
QY	949	CCACACG	CTTTGAGAG	AGCTCCACCGTCTGCTG	SAGACACCCAGG	CAATGTGGAGATCCTG	1008
DB	79361	CCACACG	CTTTGAGAG	AGCTCCACCGTCTGCTG	SAGACACCCAGG	CAATGTGGAGATCCTG	79302
QY	1009	GGCTTCT	CTGAGG	CAGAAAGGA	GAATACTTCTACAAGTATTTTCCA	CAATGTCAGAGCAG	1068
DB	79301	GGCTTCT	CTGAGG	CAGAAAGGA	GAATACTTCTACAAGTATTTTCCA	CAATGTCAGAGCAG	79242
QY	1069	GGGGG	CCAAAGTCTTCA	ATTAAGTGAAGG	CAACGAGCCTCTCTTCA	CCATGTGCTTCGTC	1128
DB	79241	GGGGG	CCAAAGTCTTCA	ATTAAGTGAAGG	CAACGAGCCTCTCTTCA	CCATGTGCTTCGTC	79182
QY	1129	CCCTCGT	GTGCTGGGTGGTGTG	TACTTGCCTCCAGCAG	CAGCTGAGAGGTGGGGGGCTG		1188
DB	79181	CCCTCGT	GTGCTGGGTGGTGTG	TACTTGCCTCCAGCAG	CAGCTGAGAGGTGGGGGGCTG		79122
QY	1189	TTGAGAC	AGACGCTCCAGG	ACCACCTGCACTG	CAGTGTACATGCTCTACCTGCTGAGTCTCATG		1248
DB	79121	TTGAGAC	AGACGCTCCAGG	ACCACCTGCACTG	CAGTGTACATGCTCTACCTGCTGAGTCTCATG		79062
QY	1249	CAACCC	AAAGCGGGG	CCCGCGCTCCAG	CCCCCACCACCCAGAGAGG	TTGTGCTCC	1308
DB	79061	CAACCC	AAAGCGGGG	CCCGCGCTCCAG	CCCCCACCACCCAGAGAGG	TTGTGCTCC	79002
QY	1309	TTGGCGG	CAGATGGGCTCTGG	AAATCAGAAATCCTATTG	AGAGCAGACCTCCGGAAG		1368
DB	79001	TTGGCGG	CAGATGGGCTCTGG	AAATCAGAAATCCTATTG	AGAGCAGACCTCCGGAAG		78942
QY	1369	CACGGC	CTTAGACGGG	GAAGACGCTCTGCTCTCTCA	AAATGAACATCTTCCAGAGGAC		1428
DB	78941	CACGGC	CTTAGACGGG	GAAGACGCTCTGCTCTCTCA	AAATGAACATCTTCCAGAGGAC		78882
QY	1429	ATCAACT	GTGAGG	TACTACAGCTTCATCTCA	CTTGAGTTTCCAGGAAATCTTTTGACGCT		1488
DB	78881	ATCAACT	GTGAGG	TACTACAGCTTCATCTCA	CTTGAGTTTCCAGGAAATCTTTTGACGCT		78822
QY	1489	ATGTACT	ATATCTCGAC	GAGGGGAGGGGCGAGCC	AGCCACACAGAGC	GTGACAGG	1548
DB	78821	ATGTACT	ATATCTCGAC	GAGGGGAGGGGCGAGCC	AGCCACACAGAGC	GTGACAGG	78762
QY	1549	CTGTTG	ACCGAGTACGCGCTTTTCTG	AAAGGAGCTTCTCTG	GCATCTCCACGACCGCTTCCTG		1608
DB	78761	CTGTTG	ACCGAGTACGCGCTTTTCTG	AAAGGAGCTTCTCTG	GCATCTCCACGACCGCTTCCTG		78702
QY	1609	TTTGG	ACTCTGAA	CGAGAGACCGAGG	ACCCTGAGAGAGTCTCTGCTG	GAAGGTC	1668
DB	78701	TTTGG	ACTCTGAA	CGAGAGACCGAGG	ACCCTGAGAGAGTCTCTGCTG	GAAGGTC	78642
QY	1669	TCGCG	CACATCAGATG	CACTGTTG	CAGTGGATCCAAAGCA	AGCTCAGAGCGACGGC	1728
DB	78641	TCGCG	CACATCAGATG	CACTGTTG	CAGTGGATCCAAAGCA	AGCTCAGAGCGACGGC	78582
QY	1729	TCCACC	CTTGACGAGG	GGCTCCTTTGAGT	TTCTTTCAGCTGCTTGTAC	GAGATCCAGAGGAG	1788
DB	78581	TCCACC	CTTGACGAGG	GGCTCCTTTGAGT	TTCTTTCAGCTGCTTGTAC	GAGATCCAGAGGAG	78522
QY	1789	GAGTTT	ATCCAGAG	CGCCCTGAGCC	ACTTCCAGGTGATCGTGGT	CAGCAACATTTGCTCC	1848
DB	78521	GAGTTT	ATCCAGAG	CGCCCTGAGCC	ACTTCCAGGTGATCGTGGT	CAGCAACATTTGCTCC	78462
QY	1849	AAGAT	GAGC	ATGGTCTCTCTG	TTCTGTCTGAAGCG	GTGACGAGCGCCAGGTCGTG	1908
DB	78461	AAGAT	GAGC	ATGGTCTCTCTG	TTCTGTCTGAAGCG	GTGACGAGCGCCAGGTCGTG	78402
QY	1909	CAC	TTGTATGCG	CCACCTACAG	CCGCGAGACCGCG	CGAGTGTCTCCGACGA	1968
DB	78401	CAC	TTGTATGCG	CCACCTACAG	CCGCGAGACCGCG	CGAGTGTCTCCGACGA	78342
QY	1969	GC	CACACG	CTGTTGGT	GCAGCT	1991	

Db 78341 GCGCACAGCTGTTGTGTCAGCT 78319
|||||

RESULT 13
AX575503

LOCUS AX575503
DEFINITION Sequence 19 from Patent WO02072630.
ACCESSION AX575503
VERSION AX575503.1 GI:27552095
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Thornton,M., Hafalia,A.J., Lu,D.A., Arvizu,C., Swarnakar,A., Lu,Y.,
Warren,B.A., Baughn,M.R., Tang,Y.T., Lee,E.A., Yao,M.G.,
Ranikumar,J.A., Khan,F., Gandhi,A.R., Ding,L., Yue,H., Gietzen,K.J.,
Walia,N.K., Thangavelu,K., Elliot,V.S. and Marquis,J.P.
TITLE Nucleic acid-associated proteins
JOURNAL Patent: WO 02072630-A 19 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..2158
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 429930CB1"
ORIGIN
Query Match 54.1%; Score 1681.8; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 1408 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 1467
Db 1 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 60
QY 1468 TTCCAGGAATCTTTTGACGTATGTACTATATCTCTGACGAGGGAGGGGGGAGGC 1527
Db 61 TTCCAGGAATCTTTTGACGTATGTACTATATCTCTGACGAGGGAGGGGGGAGGC 120
QY 1528 CCAGACCAGGACGTGACAGGCTGTTGACCGAGTACGCGTCTTTCTGAAGAGAGCTTCCTG 1587
Db 121 CCAGACCAGGACGTGACAGGCTGTTGACCGAGTACGCGTCTTTCTGAAGAGAGCTTCCTG 180
QY 1588 GCATCTACAGCGCGCTTCCTGTTTGGACTCCTCTGAAAGAGAGACCGAGCCACCTGGAG 1647
Db 181 GCATCTACAGCGCGCTTCCTGTTTGGACTCCTCTGAAAGAGAGACCGAGCCACCTGGAG 240
QY 1648 AAGAGTCTCTGCTGGAGGTCGCGCGCATCAAGATGACACCTGTTGAGTTCCTTCAGCTGC 1707
Db 241 AAGAGTCTCTGCTGGAGGTCGCGCGCATCAAGATGACACCTGTTGAGTTCCTTCAGCTGC 300
QY 1708 AGCAAGCTCAGAGCGACGCGCTCCACCTGACGAGGCTCCTTTGAGTTCCTTCAGCTGC 1767
Db 301 AGCAAGCTCAGAGCGACGCGCTCCACCTGACGAGGCTCCTTTGAGTTCCTTCAGCTGC 360
QY 1768 TTGTACGAGATCCAGAGGAGGATTTATCCAGAGCCCTGAGCCACTTCCAGGTGATC 1827
Db 361 TTGTACGAGATCCAGAGGAGGATTTATCCAGAGCCCTGAGCCACTTCCAGGTGATC 420
QY 1828 GTGGTCAGCAACATTCGCTCCAGATGGAGCAGATGGTCTCCTCGTCTGTCTGAAGCGC 1887
Db 421 GTGGTCAGCAACATTCGCTCCAGATGGAGCAGATGGTCTCCTCGTCTGTCTGAAGCGC 480
QY 1888 TGCAGGAGCGCCAGGTGCTGCACTGTATGGCGCCACCTACAGCGCGACCGGGAGAC 1947
Db 481 TGCAGGAGCGCCAGGTGCTGCACTGTATGGCGCCACCTACAGCGCGACCGGGAGAC 540
QY 1948 CGCGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTCAGCTCAGACAGAGGAGCC 2007
|||||

541 CGCGGAGGTGCTCCGAGGAGGCGACACAGCTGTGTGGTGCAGCT---ACCAGAGAGGACC 597
2008 GTTCTGCTGAGCGCTACAGTGAACATCTGCGAGCGCCCTGTGACCAATCCAAACCTG 2067
598 GTTCTGCTGAGCGCTACAGTGAACATCTGCGAGCGCCCTGTGACCAATCCAAACCTG 657
2068 ATAGAGCTGTCTCTGTACCCGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGCTCTGTCAA 2127
658 ATAGAGCTGTCTCTGTACCCGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGCTCTGTCAA 717
2128 GAGCTAGACACCCCACTGCAAACTTTCAGAACTGAGGCTGAAGAGTGGCGCATCTCC 2187
718 GAGCTAGACACCCCACTGCAAACTTTCAGAACTGAGGCTGAAGAGTGGCGCATCTCC 777
2188 AGCTCAGCGTGGAGGACCTCTCTGCGAGCTCTCATAGCAATAGAAATTTGACAAGATG 2247
778 AGCTCAGCGTGGAGGACCTCTCTGCGAGCTCTCATAGCAATAGAAATTTGACAAGATG 837
2248 GATCTCAGTGGCAACCGGCTTGGATTCAGGCAATGATGCTGCTTTCGAGGGCTCGCG 2307
838 GATCTCAGTGGCAACCGGCTTGGATTCAGGCAATGATGCTGCTTTCGAGGGCTCGCG 897
2308 CATCCCGAGTGGAGGCTGAGATGATTCAGTGGAGAGTGTGAGCTGGAGTCCGGGCT 2367
898 CATCCCGAGTGGAGGCTGAGATGATTCAGTGGAGAGTGTGAGCTGGAGTCCGGGCT 957
2368 TGTGAGGAGTGGCTTCTGTGCTCGGACCAACCACTCTGCTGAGTGGAGTCCGACTGACA 2427
958 TGTGAGGAGTGGCTTCTGTGCTCGGACCAACCACTCTGCTGAGTGGAGTCCGACTGACA 1017
2428 GGAATGCACTGGAGGATTTGGGCTGAGGTTATGTCAGGAGCTGAGGCACTCCAGTC 2487
1018 GGAATGCACTGGAGGATTTGGGCTGAGGTTATGTCAGGAGCTGAGGCACTCCAGTC 1077
2488 TGCAGACTAGCGACTTGTGCTCAAGATCTGCGGCTCACTGCTGCTGCTGAGGCTGAGGAG 2547
1078 TGCAGACTAGCGACTTGTGCTCAAGATCTGCGGCTCACTGCTGCTGCTGAGGCTGAGGAG 1137
2548 CTGCGCTCAACTCTCAGTGTGAAACAGAGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGG 2607
1138 CTGCGCTCAACTCTCAGTGTGAAACAGAGCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGG 1197
2608 CTGCGGGAACCTCGGGGCTGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2667
1198 CTGCGGGAACCTCGGGGCTGCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1257
2668 CAGACCTCGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2727
1258 CAGACCTCGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1317
2728 GTGCTCCAGGCAACCAACCTCGGGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2787
1318 GTGCTCCAGGCAACCAACCTCGGGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1377
2788 TGGGCGCTGTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2847
1378 TGGGCGCTGTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1437
2848 TGGTGGATGAGTGTGGCTCACAGCAAGGCTGTGAGATCTTTTACTTACCTCGGG 2907
1438 TGGTGGATGAGTGTGGCTCACAGCAAGGCTGTGAGATCTTTTACTTACCTCGGG 1497
2908 ATCAACAGACTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGG 2967
1498 ATCAACAGACTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGG 1557
2968 CGACTGCTTTGCAAGCGGCTGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3027
1558 CGACTGCTTTGCAAGCGGCTGAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
3028 GGGATGAGCTTGAATTAATGACCACTGAGTGGAGGCTTTCGAGTGAACCAAACT 3087
1618 GGGATGAGCTTGAATTAATGACCACTGAGTGGAGGCTTTCGAGTGAACCAAACT 1677

QY 3088 TATTGGACATTCGCTGCTGA 3108
Db 1678 TATTGGACATTCGCTGCTGA 1698

RESULT 14
AX684291
LOCUS Sequence 1 from Patent WO02052011.
DEFINITION
ACCESSION AX684291
VERSION AX684291.1 GI:29371161
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Hominidae; Homo.
REFERENCE
AUTHORS Feder, J., Ramanathan, C. and Mintier, G.
TITLE Human leucine-rich repeat containing protein, hlrbml, expressed
predominately in bone marrow
JOURNAL Patent: WO 02052011-A 1 04-JUL-2002;
Bristol-Myers Squibb Company (US)
FEATURES
source location/Qualifiers
1..4931
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1320..2669
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD86517.1"
/db_xref="GI:29371162"
/db_xref="REMTREMBL:CAD86517"
/translation="MNSQATECSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGFE
LKPSFHPDQGPWCLCWEKRPTELLNSLRKLLPELSLITTRPTALEKHLRLLEH
PRHVEILGFSEARKEYFYKFNAGAGQVENVVRNEPLFTWCFVPLVCWVCTCL
QQLEGGLRLQRTSTTTAVMLYLLSQPKAPRIQPPPNORGLCSLAADGLNQ
KILFEEDGLRGLDGDVSFAFMNFIQKDCINCEFLALTSRFLGLNDETRSHL
EKSLCWKVSPIHKMDLQWQSKAQSDGSLQQLSFFSCLYBIQEEFTQALSHF
QVIVSNIAKMEHMSVFCILKRCRQAHLVLYGATYSADGEDRARSAGATHLIVL
PERTVLDDAYSEHLAALCTNPNLIELSYRNALGSRVGLKLCQGLRHPNCKLQNR"

Query Match 52.0%; Score 1616; DB 6; Length 4931;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 5; Indels 501; Gaps 4;
QY 289 GATCCCCAGGAAACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 348
Db 1021 GATCCCCAGGAAACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 1080
QY 349 AATGCGCCCTAGGGAAATGTCAACCTCAGCCACCGGTACACCCGGCTCTCTGCTGGTG 408
Db 1081 AATGCGCCCTAGGGAAATGTCAACCTCAGCCACCGGTACACCCGGCTCTCTGCTGGTG 1140
QY 409 AAGGAGCACTCAAAACCCCATCAGGTCCAGCAGAGCTTCTGGACACAGGCGGGGACAC 468
Db 1141 AAGGAGCACTCAAAACCCCATCAGGTCCAGCAGAGCTTCTGGACACAGGCGGGGACAC 1200
QY 469 GCGAGGACCTGGGACACAGGCTAGCCCATCAGATAGAGACCTCTTTGAGCCAGAC 528
Db 1201 GCGAGGACCTGGGACACAGGCTAGCCCATCAGATAGAGACCTCTTTGAGCCAGAC 1260
QY 529 GAGGAGCGCCCGAGCCACCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 588
Db 1261 GAGGAGCGCCCGAGCCACCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
QY 589 TCCATGTGGCACACAGGATGATGTGGCTGGGCGGAGGAGCTCTTCCAGGACAG 648
Db 1314 ----- 1313
QY 649 TTGTATTATCTTCTTACATCACTGAGGAGATGAACAGAGTGGCCAGGAATGCAGC 708

Db 1314 -----AGGAGATGAACAGAGTGCACCGAATGCAGC 1346
Qy 709 ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGAGCTCATC 768
Db 1347 ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGAGCTCATC 1406
Qy 769 CGAGTTCGCGAGCGCTCTCTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCCTTTC 828
Db 1407 CGAGTTCGCGAGCGCTCTCTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCCTTTC 1466
Qy 829 CACGATCTCAGGACCTCTGTGCTCTGCTGGAGAGAAACGGCCCAACGGAGCTGCTT 888
Db 1467 CACGATCTCAGGACCTCTGTGCTCTGCTGGAGAGAAACGGCCCAACGGAGCTGCTT 1526
Qy 889 CTTAAGACCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATACACACGG 948
Db 1527 CTTAAGACCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATACACACGG 1586
Qy 949 CCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGACCCCGAGCATGTGGAGATCCTG 1008
Db 1587 CCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGACCCCGAGCATGTGGAGATCCTG 1646
Qy 1009 GGCCTTCTGAGGCAAGAAAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1068
Db 1647 GGCCTTCTGAGGCAAGAAAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1706
Qy 1069 GCGGCCAAGCTTCAATTAATGAGGAGCAACAGAGCTCTCTTACCAATGTCTTCGTC 1128
Db 1707 GCGGCCAAGCTTCAATTAATGAGGAGCAACAGAGCTCTCTTACCAATGTCTTCGTC 1766
Qy 1129 CCCCTGTGTGCTGGTGTGTACCTGCTCCAGACAGCTGAGGGTGGGGGCTG 1188
Db 1767 CCCCTGTGTGCTGGTGTGTACCTGCTCCAGACAGCTGAGGGTGGGGGCTG 1826
Qy 1189 TTGAGACAGACCTCCAGGACCACTGCGAGTGTACATCTCTACCTGCTGAGTCTGATG 1248
Db 1827 TTGAGACAGACCTCCAGGACCACTGCGAGTGTACATCTCTACCTGCTGAGTCTGATG 1886
Qy 1249 CAAACCAAGCGGGGCGCGGCTCCAGCCGCCCAACCAACAGAGAGGTTGTCTCC 1308
Db 1887 CAAACCAAGCGGGGCGCGGCTCCAGCCGCCCAACCAACAGAGAGGTTGTCTCC 1946
Qy 1309 TTGCGGCGAGTGGCTCTGGAATCAGAAATCTCTATTTGAGAGCAGGACCTCCGGAAG 1368
Db 1947 TTGCGGCGAGTGGCTCTGGAATCAGAAATCTCTATTTGAGAGCAGGACCTCCGGAAG 2006
Qy 1369 CACGGCTAGACGGGGAAGACGCTCTGCTCTTCCATCAACATGAACATCTTCCAGAGGAC 1428
Db 2007 CACGGCTAGACGGGGAAGACGCTCTGCTCTTCCATCAACATGAACATCTTCCAGAGGAC 2066
Qy 1429 ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTGAGCT 1488
Db 2067 ATCAACTGTGAG----- 2078
Qy 1489 ATGTAATAATCCTGACGAGGGGAGGGGCGGAGCCAGACCGAGCGTGACCAGG 1548
Db 2079 ----- 2078
Qy 1549 CTGTTGACCGAGTACCGGTTTTCTGAAAGAGCTTCTTGCACTCACCAGCCGCTTCCCTG 1608
Db 2079 -----AGGAGCTTCTGCACTCACCAGCCGCTTCCCTG 2111
Qy 1609 TTTGGACTCTGAAACGAGAGACACAGAGCCACTGGAGAGAGTCTCTGCTGGAGGTC 1668
Db 2112 TTTGGACTCTGAAACGAGAGACACAGAGCCACTGGAGAGAGTCTCTGCTGGAGGTC 2171
Qy 1669 TCGCCGACATCAAGATGACCTGTTGAGTGTATCCAAAGCAAGACTCAGACGCGGC 1728
Db 2172 TCGCCGACATCAAGATGACCTGTTGAGTGTATCCAAAGCAAGACTCAGACGCGGC 2231
Qy 1729 TCCACCCCTGACGAGGCTCTCTTGGAGTCTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG 1788

Db 2232 TCCACCCCTGACGAGGCTCTCTTGGAGTTCCTTTCAGCTCTTGTACGAGATCCAGGAGGAG 2291
Qy 1789 GAGTTTATCCAGCAGGCGCTCAGCACCTTCAGGATGATCGTGGTTCAGCAACATTCCTCC 1848
Db 2292 GAGTTTATCCAGCAGGCGCTCAGCACCTTCAGGATGATCGTGGTTCAGCAACATTCCTCC 2351
Qy 1849 AAGATGAGACATAGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
Db 2352 AAGATGAGACATAGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2411
Qy 1909 CACTGTTATGCGGCACTCAGCGCGGAGCGGAGAGACCGGCGAGGTGCTCCGAGGA 1968
Db 2412 CACTGTTATGCGGCACTCAGCGCGGAGCGGAGAGACCGGCGAGGTGCTCCGAGGA 2471
Qy 1969 GCGCACAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
Db 2472 GCGCACAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2528
Qy 2029 GAACATCTGGCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088
Db 2529 GAACATCTGGCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2588
Qy 2089 AATGCCCTGGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
Db 2589 AATGCCCTGGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2648
Qy 2149 AAACCTTCAGAACCTGA----- 2164
Db 2649 AAACCTTCAGAACCTGAGTAAATTTATCATATATATATATATATATATATATATATATATATATAT 2708
Qy 2165 ----- 2164
Db 2709 TTGGCCAGGTATGATGGCTCAGCGCTGTAAATTCAGCACTTTTGGAGGCGCAGATGGGGA 2768
Qy 2165 ----- 2164
Db 2769 GGATCACTTGACCCAGGAGTTCAAGACCGAGCTGCGCAACATGGTGAACCCCACTCTCTA 2828
Qy 2165 ----- 2164
Db 2829 CTAAATAATACAAATGAGCCAGGCGATGGTGGCACACGCTCTGTAGCCCGCTACTCAGG 2888
Qy 2165 -----GGCTGAAGAGGTGCC 2179
Db 2889 AGGCCAAGGCGAGGATTTGCTTCAACCCAGGAGGAGGTTGTGGCTGAAGAGGTGCC 2948
Qy 2180 GATCTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTCATAGCAATAGAAATTTGA 2239
Db 2949 GCATCTCCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTCATAGCAATAGAAATTTGA 3008
Qy 2240 CAGGATGGATCTCAGTGGCAACGGGTTGGATTTCCAGGCGATGCTGCTTGGAGG 2299
Db 3009 CAGGATGGATCTCAGTGGCAACGGGTTGGATTTCCAGGCGATGCTGCTTGGAGG 3068
Qy 2300 GCTGCGGCGATCCAGTGCAGGCTGCGAGTATTCAGTTGAGGAAGTGTCACTGCGAGT 2359
Db 3069 GCTGCGGCGATCCCAATGCGCTGCGAGTATTCAGTTGAGGAAGTGTCACTGCGAGT 3128
Qy 2360 CCGGGCTTGTGAGAGATGGCTTCTGTGCTCGGCCAACCCACATCTGTTGAGTGG 2419
Db 3129 CCGGGCTTGTGAGAGATGGCTTCTGTGCTTGGCAACCCACATCTGTTGAGTGG 3188
Qy 2420 ACTGACAGGAATGACCTGGAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTCAGGC 2479
Db 3189 ACTGACAGGAATGACCTGGAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTCAGGC 3248
Qy 2480 ACCAGTCTGACAGCTACCGAATTTTGTGCTGAAGATCTGCCGCTCACTGCTGCTGCT 2539
Db 3249 ACCAGTCTGACAGCTACCGAATTTTGTGCTGAAGATCTGCCGCTCACTGCTGCTGCT 3308
Qy 2540 GTGACAGCTGCGCTCAACTCTCAGTGTGAACCCAGAGCTGAGAGCTGGAGCTGAGCC 2599
Db 3309 GTGACAGCTGCGCTCAACTCTCAGTGTGAACCCAGAGCTGAGAGCTGGAGCTGAGCC 3368

181 ATCAACCCACCTCGGGCCAGAGGAGGCCCTGGAGGTGGCTCTCAGCACCTTTGAGCGGATA 240

ORIGIN

Db	1261	CTGATGACGCCAAGCCAGGAGCTCCAACTTCAAAGTCCACGCAACCCAGAGAGGCGCTG	1320
Qy	1303	TGCTCTCTGGCGGAGATGGGCTCTGGATCAGAAAATCCTATTTGAGGAGCAGGACCTC	1362
Db	1321	GTCTCTCTGGCTGCAGAGGCGCTCTGGATCAGAAATCTATTTGATGAACAGGATCTT	1380
Qy	1363	CGGAAGCAGCGCTAGACGGGAAGAGCTCTGCTTCTCAACATGAACATCTTCCAG	1422
Db	1381	GGGAACACGGCTAGATGGAGAGATGTCACCTTCTCACTGAACATATTCAG	1440
Qy	1423	AAGGACATCAACTGTGAGAGGTACTACAGCTTCACTTCTGAGTTTCCAGAAATCTTT	1482
Db	1441	AAGGTATCAATGTGAGAAATCTACAGCTTCACTGAGTTTCCAGAAATCTTTC	1500
Qy	1483	GCAGCTATGATATATCTTGGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1542
Db	1501	GCAGCCATGTACTGTCACTGAATGSCAGAG-----AGGCGGTG	1539
Qy	1543	ACCAGGCTGTGACGAGTACGGCTTTTCTGAAGGAGCTTCTGGCACTCACAGCCGC	1602
Db	1540	AGGAGAGCGCTGGCTGAGTATGTTTTTTCGGAAGGAATCTTCTGGCCCTCAGGTTCCAC	1599
Qy	1603	TTCTCTTTTGGACTCTCTGAACGAGGAGACCCAGAGCCACTGGAGAGAGTCTCTCTGG	1662
Db	1600	TTTCTGTTTGGCTCTCTCAACGAAGAGATGAGATGCTACCTTGAGAGGAATCTCGGCTGG	1659
Qy	1663	AAGGTCTCGCGGACATCAAGATGGAGCTGTTGCACTGGATCCAAAGCAAGCTCAGGC	1722
Db	1660	AGCATCTCCCTCAGGTGAAGGAGGAAGTGTGGCATGGATCCAAAACAAAGGCTGGGAGT	1719
Qy	1723	GACGGCTCCACCTGTCAGCAGGCGCTCTCTGGAGTCTTTCAGCTGCTGTGTACGAGATCCAG	1782
Db	1720	GAAGGCTCCACCTGTCAGCATGAGCTCCTCGAGACTACTAGCTGCTGTGTAGAGTCCAG	1779
Qy	1783	GAGGAGAGTTTATCCAGCAGGCGCTGAGCCACTTCCAGGTGATCGTGGTTCAGCAACATT	1842
Db	1780	GAGGAGGACTTTCATCCAGCAGGCGCTGAGCCACTTTCAGGTGTTGTAGTCAGAAGCATC	1839
Qy	1843	GCCTCCAAAGATGAGACATGGTCTCTCTGTTCTGTCTGAGCGCTGACGAGGCGCCAG	1902
Db	1840	TCAACAAGATGAGCACATGGTCTGCTCGTTTTTGTGCGAGGTATTGCAAGAGTACAGAA	1899
Qy	1903	GTGCTGCACTTGTATGCGCCCACTACAGCGCGGGAAGACCGCGGAGGTGCTCC	1962
Db	1900	GTGCTTCACCTTGATGGAGTGCTTTATAGTACAGGCATGGAGGACGCCACCAGAACCT	1959
Qy	1963	GCAGGAGCGCACGCTGTTGGTGCACTCAGACCAAGAGGACCGTTCTCTGGACGCC	2022
Db	1960	TCAGGAGTCCAGACTCAGTCCACATA--CTTACAGGAAGGAACATGCTGCTGTATGTC	2016
Qy	2023	TACAGTGAAATCTGGCAGCGGCGCTGTGACCAATCCAACTGATAGAGTGTCTCTG	2082
Db	2017	TACAGTGCATACCTTTTCAGCAGCTGTCTGTACCAACTCCAACTGTATCGAGCTGGCCCTTA	2076
Qy	2083	TACGGAATGCCCTGGCGAGCGGGGGTGAAGTGTCTGTCAAGGACTCAGACACCCC	2142
Db	2077	TACCGAAATGCCCTGGCGAGCGGAGGTGAAGGTGTCTGTCAAGGCTCCGACATGCC	2136
Qy	2143	AATGCAAACTTCAGAACTTGAG	2165
Db	2137	AGCTGCAAGCTGCAGAACCTGAG	2159

Search completed: August 8, 2004, 04:58:29
Job time : 11843 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 00:14:00 ; Search time 1131 Seconds
(without alignments)
11674.093 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctagcaacgcaggcag.....attggacattggctgctga 3108

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	3108	100.0	3108	6	ADE36416	Ade36416 Human PAA
2	3057	98.4	3300	6	AAL47129	Aal47129 Pyrin dom
3	2997.8	96.5	3186	6	AAL44363	Aal44363 Human PYR
4	2449.4	78.8	3306	9	ADC30316	Adc30316 Human nov
5	1705.4	54.9	1800	7	ACD03623	Novel hum
6	1681.8	54.1	2158	6	AB578719	Human cDN
7	1616	52.0	4931	6	ABL59333	Nucleotid
8	1394.8	44.9	1683	7	ADC03624	Novel hum
9	800.4	25.8	3857	4	AAD14323	Human PYR
10	800.4	25.8	3857	8	ABX93556	Huma cDNA
11	800.4	25.8	3857	8	ACD27909	Human PYR
12	639	20.6	2524	6	ADE36451	Human PAA
13	594.6	19.1	2847	4	AA301487	Human sec
14	594.6	19.1	2847	7	AB273494	Secreted
15	594.6	19.1	2847	9	ADA98038	Human sec
16	594.6	19.1	2847	9	ADC20194	Human nov
17	588.4	18.9	591	5	ADC32201	Human nov
18	477.8	15.4	506	5	AAS68757	DNA encod
19	470.4	15.1	1557	6	AAL47143	Pyrin dom
20	453.4	14.6	487	8	ACH36225	Human end
21	450.8	14.5	479	3	AAC76566	Human ORF
22	404.4	13.0	4035	6	AAL44356	Human PYR
23	402.8	13.0	3226	6	ABX97181	Human NOV

ALIGNMENTS

RESULT 1

ADE36416

ID ADE36416 standard; cDNA; 3108 BP.

XX AC ADE36416;

XX DT 29-JAN-2004 (first entry)

XX DE Human PAAD and nucleotide binding protein PAN6 cDNA.

XX KW cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
XX KW anti-allergic; antiulcer; dermatological; cerebroprotective; cardiant;
XX KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
XX KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
XX KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
XX KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
XX KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
XX KW apoptosis; NFkappaB induction; cytokine processing;
XX KW cytokine receptor; signaling caspase-mediated proteolysis;
XX KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
XX KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
XX KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
XX KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
XX KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
XX KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
XX KW ulcerative colitis; graft versus host disease; stroke; heart failure;
XX KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
XX KW cancer therapy; PAAD domain family; human; PAN6; gene; ss.

XX Homo sapiens.

XX US2003077699-A1.

XX 24-APR-2003.

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

XX 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (CHUZ/) CHU Z.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (ARIZ/) ARIZA M E.

Aad41224 Human EMB
Ada45220 Human MAT
Ada45218 Human MAT
Aal47135 Pyrin dom
Aal47131 Pyrin dom
Aal47140 Pyrin dom
Abk48628 Human MAT
Aad49018 Human MAT
Aad02764 Human NB-
Aad02764 Human NAC
Acc45443 Human NAC
Aaf83651 Human CAR
Abs55497 cDNA enco
Abs56030 cDNA enco
Acc45151 Human NAC
Acc45152 Human cas
Aad12951 Human G-P
Aad02761 Human NB-
Aad02762 Human NB-
Aad02765 Human NAC
Aal47127 Pyrin dom
Abq75801 Human MDD

Db	2581	AGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGTGTGAG	2640
Qy	2641	GGCCTCAGGCATCCACAGCTGCAAGCTCCAGACCTCGGGTTGGGCATCTGCCGCTGGGC	2700
Db	2641	GGCCTCAGGCATCCACAGCTGCAAGCTCCAGACCTCGGGTTGGGCATCTGCCGCTGGGC	2700
Qy	2701	TCTGCCGCTGTGAGGGTCTTTCTGTGGTGTCTCCAGGCCAACACACACCTCCGGGAGCTG	2760
Db	2701	TCTGCCGCTGTGAGGGTCTTTCTGTGGTGTCTCCAGGCCAACACACACCTCCGGGAGCTG	2760
Qy	2761	GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTTGTGCTGAGGGGCTGCAA	2820
Db	2761	GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTTGTGCTGAGGGGCTGCAA	2820
Qy	2821	CATCCCGCTGCAAGACTCCAGAAACTGTGSGCTGGATAGCTGTGGCCTCACAGCCCAAGGCT	2880
Db	2821	CATCCCGCTGCAAGACTCCAGAAACTGTGSGCTGGATAGCTGTGGCCTCACAGCCCAAGGCT	2880
Qy	2881	TGTGAGAATCTTTTACTTACCTCTGGGGATCAACACAGACCTTTGACCGACCTTTACCTGACC	2940
Db	2881	TGTGAGAATCTTTTACTTACCTCTGGGGATCAACACAGACCTTTGACCGACCTTTACCTGACC	2940
Qy	2941	AACAAGCCCTAGGGACACAGGTGTCGACTGCTTTGGAAGCGGTGAGCCATCTCTGGC	3000
Db	2941	AACAAGCCCTAGGGACACAGGTGTCGACTGCTTTGGAAGCGGTGAGCCATCTCTGGC	3000
Qy	3001	TGCAAACTCCGAGTCTCTGTGTTATTGGGATGAGCTGAATAAATGACCCACAGTAGG	3060
Db	3001	TGCAAACTCCGAGTCTCTGTGTTATTGGGATGAGCTGAATAAATGACCCACAGTAGG	3060
Qy	3061	TTGGCAGCGCTTCGAGTAACAAAACCTTATTTTGGACATTGGCTGCTGA	3108
Db	3061	TTGGCAGCGCTTCGAGTAACAAAACCTTATTTTGGACATTGGCTGCTGA	3108
RESULT 2			
AAAL47129			
ID	AAAL47129	standard; DNA; 3300 BP.	
AC	AAAL47129;		
XX			
XX	20-AUG-2002	(first entry)	
XX			
DE		Pyrin domain containing protein NALP3/PY5-hs coding sequence.	
XX			
KW		Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;	
KW		antiartherosclerotic; antipapillary; antibacterial; virucide;	
KW		neuroprotective; antiarthritic; antirheumatic; antiasthmatic;	
KW		nephrotropic; osteopathic; nontropic; intracellular signal transduction;	
KW		inflammation; Alzheimer's disease; infection; psoriasis; asthma;	
KW		arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;	
KW		osteoarthritis; glomerulonephritis; gene; ds.	
XX			
OS		Unidentified.	
XX			
PN		WO200240668-A2.	
XX			
PD		23-MAY-2002.	
XX			
XX		30-OCT-2001; 2001WO-EP012545.	
XX			
PR		15-NOV-2000; 2000DE-01056687.	
PR		30-NOV-2000; 2000DE-01059595.	
XX			
XX		(APOT-) APOTECH RES & DEV LTD.	
XX			
PI		Tschopp J, Martinon F;	
XX			
XX			
DR		WPI; 2002-427093/45.	
DR		P-PSDB; AAO17857.	
XX			
PT		New DNA encoding protein with pyrin domain, useful for treating diseases	

New DNA encoding protein with pyrin domain, useful for treating diseases

PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.

PS Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PVD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention

SQ Sequence 3300 BP; 726 A; 943 C; 979 G; 652 T; 0 U; 0 Other;

Query Match 98.4%; Score 3057; DB 6; Length 3300;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTACGACCGCAGCAGGACGGCTCTGTCGCTGTCCACCTACTTGGAGAACTC	60
DB	1	ATGCTACGACCGCAGCAGGACGGCTCTGTCGCTGTCCACCTACTTGGAGAACTC	60
QY	61	GAGGCTGTGGAAGTGAAGAAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGGAGAA	120
DB	61	GAGGCTGTGGAAGTGAAGAAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGGAGAA	120
QY	121	GGCAAGATCCCTGGGGAAGCATGGAGAGCGCGTCCCTGGAATGGCCGAGCTGCTC	180
DB	121	GGCAAGATCCCTGGGGAAGCATGGAGAGCGCGTCCCTGGAATGGCCGAGCTGCTC	180
QY	181	ATCACCCACTTTCGGGCGAGAGGAGCTGAGGTTGGCTCTCAGCACTTTTGACGGGATA	240
DB	181	ATCACCCACTTTCGGGCGAGAGGAGCTGAGGTTGGCTCTCAGCACTTTTGACGGGATA	240
QY	241	AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGTGGGGATCCCCAGGAA	300
DB	241	AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGTGGGGATCCCCAGGAA	300
QY	301	ACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGCAATGGCGCTTA	360
DB	301	ACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGCAATGGCGCTTA	360
QY	361	GGGGAATGTCTCAACTCAGCCACCGGTACACCGGCTCTGCTGTGTAAGAGCACTCA	420
DB	361	GGGGAATGTCTCAACTCAGCCACCGGTACACCGGCTCTGCTGTGTAAGAGCACTCA	420
QY	421	AACCCCATGAGTTCAGCAGCAGCTTCTGGAACACAGGCGGGGACACCGGAGCCGTG	480
DB	421	AACCCCATGAGTTCAGCAGCAGCTTCTGGAACACAGGCGGGGACACCGGAGCCGTG	480
QY	481	GGACACAGGCTAGCCCATCAGATAGACACCTCTTTCAGCCAGCAGAGCGGCC	540
DB	481	GGACACAGGCTAGCCCATCAGATAGACACCTCTTTCAGCCAGCAGAGCGGCC	540
QY	541	GAGCCACCGGCGACCGTGTGTCATGCAAGCGCGGAGGATAGCAAGTCCATGTGGCA	600
DB	541	GAGCCACCGGCGACCGTGTGTCATGCAAGCGCGGAGGATAGCAAGTCCATGTGGCA	600
QY	601	CACAAAGTGTGTGACTGGGCGGACGGAAAGTCTTCCAAAGGAGATTTGATATCTC	660
DB	601	CACAAAGTGTGTGACTGGGCGGACGGAAAGTCTTCCAAAGGAGATTTGATATCTC	660
QY	661	TTCTACATCACTGCAGGAGATCAACAGAGTGCACGGAATGCAGCATGCAAGCCTC	720
DB	661	TTCTACATCACTGCAGGAGATCAACAGAGTGCACGGAATGCAGCATGCAAGCCTC	720
QY	721	ATCTTACGCTGTGGCTGAGCGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG	780
DB	721	ATCTTACGCTGTGGCTGAGCGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG	780

[illegible]

Db	2941	AAACAAGCCCTTAGGGGACACAGGTGTCGACTGCTTTGCAAGCGGCTGAGCCATCTCTGGC	3000
Qy	3001	TGCAAACTCGAGTCTCTGTTATTGGGATGGACCTGTAATAAATGACCCACAGT	3057
Db	3001	TGCAAACTCGAGTCTCTGTTATTGGGATGGACCTGTAATAAATGACCCACAGT	3057
RESULT 3			
AAAL44363			
ID	AAAL44363	standard; cDNA; 3186 BP.	
XX	AAAL44363;		
AC			
XX			
DT	31-OCT-2002	(first entry)	
XX			
DE	Human PYRIN-8	cDNA sequence #2.	
XX			
KW	Human; gene; ss; gene therapy; PYRIN; stress-related response;		
KW	apoptotic response; inflammatory response; inflammatory disorder;		
KW	immune system disorder; Crohn's disease; multiple sclerosis; cancer;		
KW	leukemia; autoimmune disorder; arthritis; neurological disease;		
KW	Alzheimer's disease; Parkinson's disease; chromosomal mapping;		
KW	tissue typing; forensic biology; predictive medicine; pharmacogenomics;		
KW	transcription profiling; PYRIN-8.		
XX			
OS	Homo sapiens.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3186	
FT		/*cag= a	
FT		/product= "Human PYRIN-8 protein #2"	
XX			
XX			
PN	WO200261049-A2.		
XX			
PD	08-AUG-2002.		
XX			
PF	31-JAN-2002; 2002WO-US002967.		
XX			
PR	31-JAN-2001; 2001US-0265231P.		
PR	10-SEP-2001; 2001US-0318645P.		
XX			
FA	(MILL-) MILLENNIUM PHARM INC.		
PA	(AMHP) WYETH.		
XX			
FI	Bertin J, Wang W, Blatcher M;		
XX			
DR	WPI; 2002-627477/67.		
XX	P-PSDE; AAO15590.		
DR			
XX			
PT	New PYRIN polypeptides and nucleic acids useful for modulating and		
PT	diagnosing stress-related, apoptotic and inflammatory responses, or for		
PT	treating inflammatory and immune system disorders, cancers, or		
PT	neurological diseases.		
XX			
PS	Claim 4; Fig 8; 167pp; English.		
XX			
CC	The invention comprises the amino acid and coding sequences of human		
CC	PYRIN proteins. The PYRIN protein and DNA sequences of the invention are		
CC	useful for modulating and diagnosing stress-related, apoptotic and		
CC	inflammatory responses. The PYRIN protein and DNA sequences are useful		
CC	for treating: inflammatory disorders and immune system disorders (e.g.		
CC	Crohn's disease, reactive arthritis, multiple sclerosis, contact		
CC	dermatitis, psoriasis, graft rejection, allergies, viral infections and		
CC	bacterial infections); cancer (e.g. leukaemia); autoimmune disorders		
CC	(e.g. systemic lupus erythematosus and arthritis); and neurological		
CC	diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN		
CC	protein and DNA sequences may also be used in screening assays, detection		
CC	assays (e.g. chromosomal mapping, tissue typing or forensic biology),		
CC	predictive medicine (e.g. diagnostic assays, clinical trials and		
CC	pharmacogenomics) and transcription profiling. The present DNA sequence		
CC	encodes a human PYRIN-8 protein		
XX			
Sequence	3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other;		

2098	Db	GCCTACAGTGAACATCTGGCAGCGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCT	2157
2080	Qy	CTGTACCGAAATGCCCTTGGGCAGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC	2139
2158	Db	CTGTACCGAAATGCCCTTGGGCAGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC	2217
2140	Qy	CCNACTGC AAACTTCA GAAACCTGAGGCTGAAGAGGTGCCGCACTCTCCAGCTCAGCCTGC	2199
2218	Db	CCCAACTGC AAACTTCA GAAACCTGAGGCTGAAGAGGTGCCGCACTCTCCAGCTCAGCCTGC	2277
2200	Qy	GAGGACCTCTCGACGCTCATAGCCAAATAAGAAATTTGACAGGATGGATCTCACTGCGC	2259
2278	Db	GAGGACCTCTCTCGACGCTCATAGCCAAATAAGAAATTTGACAGGATGGATCTCACTGCGC	2337
2260	Qy	AACGGCGTTGGATTTCCAGGCATGATGCTCTTTGGAGGGCCCTGCGGCATCCCAAGTGC	2319
2338	Db	AACGGCGTTGGATTTCCAGGCATGATGCTCTTTGGAGGGCCCTGCGGCATCCCAAGTGC	2397
2320	Qy	AGGCTGCAGATGATTCAGTTGAGAAAGTGTCACTGGAGTCCGGGCTTTGTCAAGAGATG	2379
2398	Db	AGGCTGCAGATGATTCAGTTGAGAAAGTGTCACTGGAGTCCGGGCTTTGTCAAGAGATG	2457
2380	Qy	GCTTCTGCTCGCCACCAACCCACATCTGGTTGAGTTGAGTTGACCTGCACAGGAATGCACTG	2439
2458	Db	GCTTCTGCTCGCCACCAACCCACATCTGGTTGAGTTGAGTTGACCTGCACAGGAATGCACTG	2517
2440	Qy	GAGGATTTGGGCTGAGGTTACTATGCAGGGACTGAGGCACCCAGTCTGTGCAGACTACCG	2499
2518	Db	GAGGATTTGGGCTGAGGTTACTATGCAGGGACTGAGGCACCCAGTCTGTGCAGACTACCG	2577
2500	Qy	ACTTTGTGGCTGAAGATCTGCCGCTCACTGTCTGCTGCTGTGACGAGCTGGGCTCAACT	2559
2578	Db	ACTTTGTGGCTGAAGATCTGCCGCTCACTGTCTGCTGCTGTGACGAGCTGGGCTCAACT	2637
2560	Qy	CTCAGTGTGAACAGAGCCTGAGAGACTGGACCTGAGCCTGATGAGCTGGGGGACCTC	2619
2638	Db	CTCAGTGTGAACAGAGCCTGAGAGACTGGACCTGAGCCTGATGAGCTGGGGGACCTC	2697
2620	Qy	GGGGTCTGCTGTGTGAGGGCCCTCAGGCATCCACGTGCAAGCTCCAGACCCCTGGCG	2679
2698	Db	GGGGTCTGCTGTGTGAGGGCCCTCAGGCATCCACGTGCAAGCTCCAGACCCCTGGCG	2757
2680	Qy	TTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTGTGTCTCCAGGCC	2739
2758	Db	TTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTGTGTCTCCAGGCC	2817
2740	Qy	AACCAAACTCCGGAGCTGGACTTGAATTCACGACTGGGAGACTGGGGCTGTGG	2799
2818	Db	AACCAAACTCCGGAGCTGGACTTGAATTCACGACTGGGAGACTGGGGCTGTGG	2877
2800	Qy	TTGCTGGCTGAGGGCTCAACATCCCGCTGCAGACTCCAGAAATCTGTGGCTGGATAGC	2859
2878	Db	TTGCTGGCTGAGGGCTCAACATCCCGCTGCAGACTCCAGAAATCTGTGGCTGGATAGC	2937
2860	Qy	TGTGGCTCACAGCCAGGCTGTGAAATCTTTACTTCAACCTGGGGATCAACAGACC	2919
2938	Db	TGTGGCTCACAGCCAGGCTGTGAAATCTTTACTTCAACCTGGGGATCAACAGACC	2997
2920	Qy	TTGACCGACTTTACTGTACCAACAGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGC	2979
2998	Db	TTGACCGACTTTACTGTACCAACAGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGC	3057
2980	Qy	AACGGCTGAGCCATCCTGTGCTCAAACTCCGAGTCTCTGTGTTATTTGGGATGGACCTG	3039
3058	Db	AACGGCTGAGCCATCCTGTGCTCAAACTCCGAGTCTCTGTGTTATTTGGGATGGACCTG	3117
3040	Qy	AATAAAATACCCACAGTAGTTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACATT	3099
3118	Db	AATAAAATACCCACAGTAGTTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACATT	3177
3100	Qy	GGCTGTGTGA	3108

Db	3178	GGCTGCTGA	3186
RESULT 4			
AD3C30316			
ID	AD3C30316	standard; cDNA; 3306 BP.	
XX	AC	AD3C30316;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	Human novel cDNA sequence, SEQ ID NO:398.	
XX	XX	Human; diagnostic; drug screening; forensics; gene mapping;	
KW	KW	biodiversity assessment; Parkinson's disease; Alzheimer's disease;	
KW	KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;	
KW	KW	ulcers; osteoporosis; autoimmune disease; cancer;	
KW	KW	molecular weight marker; food supplement; antiparkinsonian; nootropic;	
KW	KW	neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;	
KW	KW	antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;	
KW	KW	gene therapy; chromosome 19; gene; ss.	
XX	XX		
OS	OS	Homo sapiens.	
XX	XX		
PN	PN	WO2003029271-A2.	
XX	XX		
PD	PD	10-APR-2003.	
XX	XX		
PF	PF	24-SEP-2002; 2002WO-US030474.	
XX	XX		
PR	PR	24-SEP-2001; 2001US-0324631P.	
XX	XX	(HYSE-) HYSEQ INC.	
PA	PA		
PI	PI	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;	
PI	PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;	
PI	PI	Halley-Vicente D, Drmanac RT;	
XX	XX		
DR	DR	WPI: 2003-371981/35.	
DR	DR	P-PSDB; AD3C31287.	
XX	XX		
PT	PT	New polynucleotide and polypeptide useful for diagnosing, preventing or	
PT	PT	treating conditions such as neurodegenerative diseases, anemias, platelet	
PT	PT	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or	
PT	PT	cancer.	
XX	XX		
PS	PS	Claim 1; SEQ ID NO 398; 1185pp; English.	
XX	XX		
CC	CC	The invention relates to 971 novel human cDNA sequences (AD3C29919-	
CC	CC	AD3C30889) and the polypeptides they encode (AD3C30890-AD3C31860). The	
CC	CC	invention also relates to nucleic acid sequences over 99% identical with	
CC	CC	the novel human cDNAs. The invention additionally encompasses expression	
CC	CC	vectors and host cells comprising a nucleic acid of the invention; an antibody	
CC	CC	recombinant production of a polypeptide of the invention; a method of detecting	
CC	CC	against a polypeptide of the invention; a method of detecting	
CC	CC	polynucleotides or polypeptides of the invention; and methods of	
CC	CC	identifying a compound which binds to a polypeptide of the invention. The	
CC	CC	invention further discloses methods of preventing, treating or	
CC	CC	ameliorating a medical condition; kits comprising polynucleotide probes	
CC	CC	and/or monoclonal antibodies for carrying out the methods of the	
CC	CC	invention; methods for the identification of compounds that modulate the	
CC	CC	expression or activity of the polynucleotide and/or polypeptide; and 767	
CC	CC	contig sequences corresponding to the cDNA sequences of the invention	
CC	CC	(AD3C31861-AD3C32627) and the polypeptides encoded by the contigs (AD3C3262	
CC	CC	-AD3C33394). The nucleic acids and polypeptides of the invention are	
CC	CC	useful in diagnostics, drug screening, forensics, gene mapping, in the	
CC	CC	identification of mutations responsible for genetic disorders or other	
CC	CC	traits, for assessing biodiversity, and in producing many other types of	
CC	CC	data and products dependent on DNA and amino acid sequences. They are	
CC	CC	also used for treating diseases such as Parkinson's disease, Alzheimer's	
CC	CC	disease and other neurodegenerative diseases, anaemia, platelet	
CC	CC	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or	
CC	CC	cancer. The nucleic acids may also be used as hybridization probes or	

CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC fp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3306 BP; 718 A; 958 C; 939 G; 691 T; 0 U; 0 Other;

Query Match 78.8%; Score 2449.4; DB 9; Length 3306;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 2644; Conservative 0; Mismatches 1; Indels 174; Gaps 2;

Qy	290	ATCCCCAGGAAACCTACAGGAGTATGTCGACAGAAATTCGGCTCATGGAAGACCGCA	349
Db	498	ATCCCCAGGAAACCTACAGGAGTATGTCGACAGAAATTCGGCTCATGGAAGACCGCA	557
Qy	350	ATGCGCGCTAGGGAATGTCTAACTCAGCAGCAGCTTCTGGACACAGCGCGGGACACG	409
Db	558	ATGCGCGCTAGGGAATGTCTAACTCAGCAGCAGCTTCTGGACACAGCGCGGGACACG	617
Qy	410	AGGAGCACTCAAAACCCCATGCTCCAGCAGCAGCTTCTGGACACAGCGCGGGACACG	469
Db	618	AGGAGCACTCAAAACCCCATGCTCCAGCAGCAGCTTCTGGACACAGCGCGGGACACG	677
Qy	470	CGAGGACCGTGGGACACAGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGCG	529
Db	678	CGAGGACCGTGGGACACAGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGCG	737
Qy	530	AGGAGCGCCCGGACCGCAGCTGCTGAGTGGCGGACGGAGCTCTTCAAGCGAGAT	589
Db	738	AGGAGCGCCCGGACCGCAGCTGCTGAGTGGCGGACGGAGCTCTTCAAGCGAGAT	797
Qy	590	CCATGCTGGACACAAAGTGTGCTGCTGAGTGGCGGACGGAGCTCTTCAAGCGAGAT	649
Db	798	CCATGCTGGACACAAAGTGTGCTGAGTGGCGGACGGAGCTCTTCAAGCGAGAT	857
Qy	650	TTGATTATCTTCTATCACTCACTGCGGGAGATGAACAGAGTGCCACGGAATGCAGCA	709
Db	858	TTGATTATCTTCTATCACTCACTGCGGGAGATGAACAGAGTGCCACGGAATGCAGCA	917
Qy	710	TGCAAGACCTCATCTTCACTGCTGCTGAGTGGCGGACGGCGCTCTCCAGAGCTCATCC	769
Db	918	TGCAAGACCTCATCTTCACTGCTGCTGAGTGGCGGACGGCGCTCTCCAGAGCTCATCC	977
Qy	770	GAGTTCCCGAGCGCTCTCTTTCATCATCGACGCTTCGATGAGTCAAGCTTCTTTCC	829
Db	978	GAGTTCCCGAGCGCTCTCTTTCATCATCGACGCTTCGATGAGTCAAGCTTCTTTCC	1037
Qy	830	ACGATCTCAGGACCTCTGCTGCTGCTGAGTGGAGGAGAAACCGCCACCGAGTGTTC	889
Db	1038	ACGATCTCAGGACCTCTGCTGCTGCTGAGTGGAGGAGAAACCGCCACCGAGTGTTC	1097
Qy	890	TTAAGACCTTAATTCGAAGAAGCTGCTCCCTGAGTCTCTTCTCATCACACACGCG	949
Db	1098	TTAAGACCTTAATTCGAAGAAGCTGCTCCCTGAGTCTCTTCTCATCACACACGCG	1157
Qy	950	CCAGCGCTTTGGAGAACTCCACCGTCTGCTGAGCAGCCCGGAGCATGTGAGATCCTGG	1009
Db	1158	CCAGCGCTTTGGAGAACTCCACCGTCTGCTGAGCAGCCCGGAGCATGTGAGATCCTGG	1217
Qy	1010	GCTTCTCTGAGGCAAGAAAGAAATCTTCTCAAGATTTTCAAGTTCAGAGCAGG	1069
Db	1218	GCTTCTCTGAGGCAAGAAAGAAATCTTCTCAAGATTTTCAAGTTCAGAGCAGG	1277
Qy	1070	CGGCGCAAGTCTCAATTCAGTGGGACACAGAGCTCTTCCACATGCTTCTCTCC	1129
Db	1278	CGGCGCAAGTCTCAATTCAGTGGGACACAGAGCTCTTCCACATGCTTCTCTCC	1337
Qy	1130	CCCTGCTGCTGGTGGTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1189
Db	1338	CCCTGCTGCTGGTGGTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1397

Qy	1190	TGAGACAGAGCTCCAGGACCAACCACTCAGTGTACATGCTTACCTCTGAGTCTGATGC	1249
Db	1398	TGAGACAGAGCTCCAGGACCAACCACTCAGTGTACATGCTTACCTCTGAGTCTGATGC	1457
Qy	1250	AACCCAAAGCGGGGGCCCCCGGCTCCAGCCCCCACCACCAACAGAGAGGTTGCTCTCT	1309
Db	1458	AACCCAAAGCGGGGGCCCCCGGCTCCAGCCCCCACCACCAACAGAGAGGTTGCTCTCT	1517
Qy	1310	TGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTAGGAGCAGACCTCCGGAAGC	1369
Db	1518	TGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTAGGAGCAGACCTCCGGAAGC	1577
Qy	1370	ACGCGCTAGACGGGGAAGAAGTCTCTGCTTCTTCAATGAAATCTTCCAGAAAGACA	1429
Db	1578	ACGCGCTAGACGGGGAAGAAGTCTCTGCTTCTTCAATGAAATCTTCCAGAAAGACA	1637
Qy	1430	TCAACTGTGAGAGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTTCAGCTA	1489
Db	1638	TCAACTGTGAGAGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTTCAGCTA	1697
Qy	1490	TGTACTATATCTCGACGAGGGGAGGGCGGGGAGGCCAGAGACCGTGAACAGGC	1549
Db	1698	TGTACTATATCTCGACGAGGGGAGGGCGGGGAGGCCAGAGACCGTGAACAGGC	1757
Qy	1550	TGTTGACCGAGTACCGCTTTTCTGAAAGAGTCTTCTGGCACTCAACAGCGCTTCTCTGT	1609
Db	1758	TGTTGACCGAGTACCGCTTTTCTGAAAGAGTCTTCTGGCACTCAACAGCGCTTCTCTGT	1817
Qy	1610	TTGGAGCTCTGAGGAGGACAGGAGCCACCTGAGAGAGAGTCTCTGCTGAGAGTCT	1669
Db	1818	TTGGAGCTCTGAGGAGGACAGGAGCCACCTGAGAGAGAGTCTCTGCTGAGAGTCT	1877
Qy	1670	CGCCGCACTCAAGATGAGCTGTTGAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGCT	1729
Db	1878	CGCCGCACTCAAGATGAGCTGTTGAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGCT	1937
Qy	1730	CCACCTTGACAGCGGCTCTTGGAGTCTTTCAGTCTGTTAGAGATCCAGAGGAGG	1789
Db	1938	CCACCTTGACAGCGGCTCTTGGAGTCTTTCAGTCTGTTAGAGATCCAGAGGAGG	1997
Qy	1790	AGTTTATCCAGCGGCTCAGCCACTTCCAGTGTGCTGCTGAGAGAGTCTGCTGAGAGTCT	1849
Db	1998	AGTTTATCCAGCGGCTCAGCCACTTCCAGTGTGCTGCTGAGAGAGTCTGCTGAGAGTCT	2057
Qy	1850	AGATGAGCAGATGGTCTCTCTGTTCTGTCTGAAGCGCTGAGAGGCGCCAGAGTGTGC	1909
Db	2058	AGATGAGCAGATGGTCTCTCTGTTCTGTCTGAAGCGCTGAGAGGCGCCAGAGTGTGC	2117
Qy	1910	ACTTGTATGGCGCCACTTACAGCGCGGAGCGGGGAAGACCGCGCGAGTGTCTCCGAGGAG	1969
Db	2118	ACTTGTATGGCGCCACTTACAGCGCGGAGCGGGGAAGACCGCGCGAGTGTCTCCGAGGAG	2177
Qy	1970	CGCACAGCTGTTGGTGCAGCTCAGCAGAGGAGCGGTTCTGCTGAGCGCTTACAGTGTG	2029
Db	2178	CGCACAGCTGTTGGTGCAGCTCAGCAGAGGAGCGGTTCTGCTGAGCGCTTACAGTGTG	2234
Qy	2030	AACTCTGGAGCGGCGCTGTGCAACCAATCAAACTGATAGAGTCTGCTGTACCGAA	2089
Db	2235	AACTCTGGAGCGGCGCTGTGCAACCAATCAAACTGATAGAGTCTGCTGTACCGAA	2294
Qy	2090	ATGCGCTGGGAGCGCGGGGGTGAAGTGTCTGTCAAGAGCTCAGACACCCCAACTGCA	2149
Db	2295	ATGCGCTGGGAGCGCGGGGGTGAAGTGTCTGTCAAGAGCTCAGACACCCCAACTGCA	2354
Qy	2150	AACCTTCAGAACTCAGAGCTGAAGAGTCCGCTCTCCAGTCTGCTGAGAGCTCTCT	2209
Db	2355	AACCTTCAGAACTCAGAGCTGAAGAGTCCGCTCTCCAGTCTGCTGAGAGCTCTCT	2414
Qy	2210	CTGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGAGTCTCAGTGGCAACCGCGTTG	2269
Db	2415	CTGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGAGTCTCAGTGGCAACCGCGTTG	2474

QY 649 TTGATTATCTCTTACATCAACTGCAGGGAGATGAACACAGATGCCACGGAATGCAGC 708
Db 1314 -----AGGGAGATGAACACAGATGCCACGGAATGCAGC 1346
QY 709 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCGACGCGGCCCTCTCCAGGAGCTCATC 768
Db 1347 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCGACGCGGCCCTCTCCAGGAGCTCATC 1406
QY 769 CGAGTTCCGAGCGCTCTCTTTTCATCATCGACGCTTCGATGAGCTCAAGCTCTCTTTC 828
Db 1407 CGAGTTCCGAGCGCTCTCTTTTCATCATCGACGCTTCGATGAGCTCAAGCTCTCTTTC 1466
QY 829 CACGATCTCAGGACCTCTGTGCTCTCTGCTGGAGGAGAAACGCGCCACGAGCTGTCT 888
Db 1467 CACGATCTCAGGACCTCTGTGCTCTCTGCTGGAGGAGAAACGCGCCACGAGCTGTCT 1526
QY 889 CTTAAGACCTTAATTTCGAAGAGCTGCTCCCTGAGCTATCTTTGCTGCTATCAACACGG 948
Db 1527 CTTAAGACCTTAATTTCGAAGAGCTGCTCCCTGAGCTATCTTTGCTGCTATCAACACGG 1586
QY 949 CCCACGCTTTGGAGAAGCTCCACCGTCTGCTGGAGACACCCAGGACATGTGGAGATCCTG 1008
Db 1587 CCCACGCTTTGGAGAAGCTCCACCGTCTGCTGGAGACACCCAGGACATGTGGAGATCCTG 1646
QY 1009 GCCTTCTCTGAGGCGAAGGAAGAAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1068
Db 1647 GCCTTCTCTGAGGCGAAGGAAGAAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1706
QY 1069 GGGGCCAAGTCTCAANTTACGTAGGAGACACGAGCCCTCTTTCACCATGTGCTTCGTC 1128
Db 1707 GGGGCCAAGTCTTCAANTTACGTAGGAGACACGAGCCCTCTTTCACCATGTGCTTCGTC 1766
QY 1129 CCCCTGTGTGCTGGTGTGATCTGCTGCTCCAGCAGCAGCTCGAGGCTGGGGGGCTG 1188
Db 1767 CCCCTGTGTGCTGGTGTGATCTGCTGCTCCAGCAGCAGCTCGAGGCTGGGGGGCTG 1826
QY 1189 TTGACAGACGCTCAGGACACCACTGCACTGTATCTCTACCTGCTGAGTCTGATG 1248
Db 1827 TTGACAGACGCTCAGGACACCACTGCACTGTATCTCTACCTGCTGAGTCTGATG 1886
QY 1249 CAACCAAGCCGGGGCCCGCGCTCCAGCCGCCCAACCAAGAGGGTTGTGCTCC 1308
Db 1887 CAACCAAGCCGGGGCCCGCGCTCCAGCCGCCCAACCAAGAGGGTTGTGCTCC 1946
QY 1309 TTGGCGGAGATGGCTCTGGAATCAGAAAATCTTATTTGAGAGCAGACCTCCGGAAG 1368
Db 1947 TTGGCGGAGATGGCTCTGGAATCAGAAAATCTTATTTGAGAGCAGACCTCCGGAAG 2006
QY 1369 CACGCTCTAGACGGGGAAGAGCTCTGCTCTTCTCAACATGAACATCTTCCAGAAGGAC 1428
Db 2007 CACGCTCTAGACGGGGAAGAGCTCTGCTCTTCTCAACATGAACATCTTCCAGAAGGAC 2066
QY 1429 ATCACTGTGAGAGTACTACAGCTTCATCCTGAGTTTCCAGGAATTTCTTGCAGCT 1488
Db 2067 ATCACTGTGAG----- 2078
QY 1489 ATGTACTATATCTTGAACGAGGGGAGGCGGGCAGGCCACAGACGAGACGTGACCAGG 1548
Db 2079 ----- 2078
QY 1549 CTGTTGACCGAGTACGCGTTTCTGAAAGAGCTTCTTGCAAGAGCTCACCAGCGCTTCTG 1608
Db 2079 -----AGGAGCTTCTTGGCACTCACCAGCGCTTCTG 2111
QY 1609 TTTGACTCTTGAACGAGGAGACGAGGAGCCACTGGAGAGAGTCTCTGCTGGAAGGTC 1668
Db 2112 TTTGACTCTTGAACGAGGAGACGAGGAGCCACTGGAGAGAGTCTCTGCTGGAAGGTC 2171
QY 1669 TCGCGGCACATCAAGATGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGACGGC 1728
Db 2172 TCGCGGCACATCAAGATGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGACGGC 2231
QY 1729 TCCACCTGCGACAGGGCTCTTGGAGTTCTTCAGCTGCTTGTACGAGATCCAGGAGGAG 1788

Db 2232 TCCACCTGCGACAGGGCTCTTGGAGTTCTTCAGCTGCTTACGAGATCCAGGAGGAG 2291
QY 1789 GAGTTTATCCAGCAGCGCTTACGACCTTTCAGGTTGATCGTGGTCAAGCAATTCCTCC 1848
Db 2292 GAGTTTATCCAGCAGCGCTTACGACCTTTCAGGTTGATCGTGGTCAAGCAATTCCTCC 2351
QY 1849 AAGATGGAGCACATGCTCTCTGCTTCTGCTGAAGCGCTGAGAGCGCCAGGTCGTCG 1908
Db 2352 AAGATGGAGCACATGCTCTCTGCTTCTGCTGAAGCGCTGAGAGCGCCAGGTCGTCG 2411
QY 1909 CACTTGTATGGCCACCTTACAGCGCGGACGGGGAAGACCGCGCAGGTTGCTCCGAGGA 1968
Db 2412 CACTTGTATGGCCACCTTACAGCGCGGACGGGGAAGACCGCGCAGGTTGCTCCGAGGA 2471
QY 1969 GCGCACACGCTGTGCTGCTGAGCTCAGACGAGAGGACGGTCTCTGCTGAGCGCCTACAGT 2028
Db 2472 GCGCACACGCTGTGCTGCTGAGCT---ACCAGAGAGCGCTTCTGCTGAGCGCCTACAGT 2528
QY 2029 GAAACATCTGGCAGCGCCCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA 2088
Db 2529 GAAACATCTGGCAGCGCCCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA 2588
QY 2089 AATGCCCTGGCAGCGCGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCAACTGC 2148
Db 2589 AATGCCCTGGCAGCGCGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCAACTGC 2648
QY 2149 AAACCTCAGAACCTGA----- 2164
Db 2649 AAACCTCAGAACCTGAGGTAAATTTATCATATATAACATGATATTTTGAATAAATATA 2708
QY 2165 ----- 2164
Db 2709 TTGGCCAGTATGATGGCTCAGCGCTGTAATTCAGACACTTTGGAGGCGCCAGATGGGA 2768
QY 2165 ----- 2164
Db 2769 GATCATTGACCCAGGAGTTCAAGACACCGCTGCGCAACATGGTGAACCCCACTCTTA 2828
QY 2165 ----- 2164
Db 2829 CTAAAAATACCAAAATGAGCCAGGCGATGCTGGCACACGCTCTGTAAGCCAGCTACTCAGG 2888
QY 2165 -----GGCTGAAGAGTGCC 2179
Db 2889 AGGCCAAGCAGGAGGATGCTTCAACCCAGGAGCAGAGTTGTGGCTGAAGAGTGCC 2948
QY 2180 GCATCTCCAGCTCAGCCTCGGAGGACCTCTCTGAGCTCTCATAGCCCAATAAGAAATTTGA 2239
Db 2949 GCATCTCCAGCTCAGCCTCGGAGGACCTCTCTGAGCTCTCATAGCCCAATAAGAAATTTGA 3008
QY 2240 CAAGATGGATCTCAGTGGCAACCGCGTTGATTCAGGAGGATGATGCTGCTTTGGAGG 2299
Db 3009 CAAGATGGATCTCAGTGGCAACCGCGTTGATTCAGGAGGATGATGCTGCTTTGGAGG 3068
QY 2300 GCCTGCGCATCCCGCAGTGCGAGCTGAGATGATTCAGTTGAGGAGTGTGAGCTGGAGT 2359
Db 3069 GCCTGCGCATCCCGCAGTGCGAGCTGAGATGATTCAGTTGAGGAGTGTGAGCTGGAGT 3128
QY 2360 CCGGGCTTGTCAAGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGTTGAGTTGG 2419
Db 3129 CCGGGCTTGTCAAGAGATGGCTTCTGTGCTTGGCACCACCCACATCTGTTGAGTTGG 3188
QY 2420 ACCTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCGAGGAGCTGAGGC 2479
Db 3189 ACCTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCGAGGAGCTGAGGC 3248
QY 2480 ACCAGCTCTGAGACTACGAGCTTGTGGCTGAGATCTGCGCCTCACTGCTGCTGCT 2539
Db 3249 ACCAGCTCTGAGACTACGAGCTTGTGGCTGAGATCTGCGCCTCACTGCTGCTGCT 3308
QY 2540 GTGACGAGCTGCGCTCAACTCTCAGTGTGAACCCAGAGGCTGAGAGCTGAGACCTGAGCC 2599

Db 542 CCTGGAGAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAGATGAGCTGTGTCAGTG 601
QY 1701 GATCCAAAGCAAGCTCAGAGCGAGCGCTCCACCTCGCAGCGAGGCTCCTCTGAGTTCCT 1760
Db 602 GATCCAAAGCAAGCTCAGAGCGAGCGCTCCACCTCGCAGCGAGGCTCCTCTGAGTTCCT 661
QY 1761 CAGCTCTTTGTAGAGATCCAGGAGGAGGAGTTTATCCAGCAGCGCTCAGCCACATCCCA 1820
Db 662 CAGCTCTTTGTAGAGATCCAGGAGGAGGAGTTTATCCAGCAGCGCTCAGCCACATCCCA 721
QY 1821 GGTGATCGTGTGAGCAATTCCTCCAGATGAGAGCAGATGCTCTCTGCTGTCTGCT 1880
Db 722 GGTGATCGTGTGAGCAATTCCTCCAGATGAGAGCAGATGCTCTCTGCTGTCTGCT 781
QY 1881 GAAGCGCTCAGAGCGCGCAGGTGCTGCACTGTATGCGCCACCTACAGCGCGAGCG 1940
Db 782 GAAGCGCTCAGAGCGCGCAGGTGCTGCACTGTATGCGCCACCTACAGCGCGAGCG 841
QY 1941 GGAAGACCGCGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACCAGA 2000
Db 842 GGAAGACCGCGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACCAGA 901
QY 2001 GAGGACCGTCTGCTGGAGCGCTCAGTGAACATCTGGCAGCGCGCTGTGACCAATCC 2060
Db 902 GAGGACCGTCTGCTGGAGCGCTCAGTGAACATCTGGCAGCGCGCTGTGACCAATCC 961
QY 2061 AAACCTGATAGAGTCTCTCTGTACCGAATGCTGCGCAGCGCGGGGTGAAGTCT 2120
Db 962 AAACCTGATAGAGTCTCTCTGTACCGAATGCTGCGCAGCGCGGGGTGAAGTCT 1021
QY 2121 CTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAACTGAGGCTGAAGAGTGGCG 2180
Db 1022 CTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAACTGAGGCTGAAGAGTGGCG 1081
QY 2181 CATCTCCAGCTCAGCTCGAGGAGCTCTCTGAGCTCTCATAGCAATTAAGATTTGAC 2240
Db 1082 CATCTCCAGCTCAGCTCGAGGAGCTCTCTGAGCTCTCATAGCAATTAAGATTTGAC 1141
QY 2241 AAGGATGATCTCAGTGGCAACGGGGTGGATTCCTCAGGATGATGCTGTTGGAGGG 2300
Db 1142 AAGGATGATCTCAGTGGCAACGGGGTGGATTCCTCAGGATGATGCTGTTGGAGGG 1201
QY 2301 CTTGGGATCCCACTGAGCTGAGATGATTCAGTTGAGGAAGTGTGAGCTGAGTGC 2360
Db 1202 CTTGGGATCCCACTGAGCTGAGATGATTCAGTTGAGGAAGTGTGAGCTGAGTGC 1237
QY 2361 CGGGCTTGTAGGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGTTGAGTTGA 2420
Db 1238 ----- 1237
QY 2421 CTGACAGGAAATGCACTGAGGATTTGGGCTGAGGTTACTATGCCAGGAGTGAAGCA 2480
Db 1238 ----- 1237
QY 2481 CCCAGCTGCGAGCTACGAGCTTGTGGTGAAGATCTGCCCTCACTGCTGCTGCTG 2540
Db 1238 -----GCTGAAGATCTGCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1270
QY 2541 TGACAGCTGGGCTCACTCTCAGTGTGTAACAGAGCTGAGAGCTGAGCTGAGCTGAGCT 2600
Db 1271 TGACAGCTGGGCTCACTCTCAGTGTGTAACAGAGCTGAGAGCTGAGCTGAGCTGAGCT 1330
QY 2601 GAATGAGCTGGGAGCTCGGGGTGCTGCTGTGTGAGGAGGCTCAGGATCCCACTG 2660
Db 1331 GAATGAGCTGGGAGCTCGGGGTGCTGCTGTGTGAGGAGGCTCAGGATCCCACTG 1390
QY 2661 CAAGCTCCAGACCTCGGGTGTGGGATCTGCCGCTGGGCTCTGCGGCTGTGAGGCTCT 2720
Db 1391 CAAGCTCCAGACCTCGGGTGTGGGATCTGCCGCTGGGCTCTGCGGCTGTGAGGCTCT 1450
QY 2721 TTTCTGTGTCTCAGGCGCAACCACTCTCGGAGCTGAGTTCAGGACTTCAACGACT 2780
|||||

Db 1451 TTCTGTGTGCTCCAGGCGCAACCAACCTCCGGAGCTGAGTTCAGGACT 1510
QY 2781 GGGAGACTGGGCTGTGTTGCTGCTGCTGAGGGCTGCAACATCCCGCTGCGAGCTCCA 2840
Db 1511 GGGAGACTGGGCTGTGTTGCTGCTGCTGAGGGCTGCAACATCCCGCTGCGAGCTCCA 1570
QY 2841 GAAACTGTGG 2850
Db 1571 GAAACTGTGG 1580
RESULT 9
AAD14323
ID AAD14323 standard; cDNA; 3857 BP.
AC AAD14323;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human PYRIN 1 (PYRIN-1) cDNA.
XX
KW Human; nucleotide binding site; NBS; PYRIN-1; apoptotic; cytostatic;
dermatological; vasotropic; melanoma; gene therapy; ophthalmological;
cancer; leukaemia; carcinoma; arthritis; viral infection; allergy;
autoimmune disease; systemic lupus erythematosus; SLE; neutropenic; asthma;
glomerulonephritis; neurological disorder; Alzheimer's disease; ALS;
amyotrophic lateral sclerosis; Parkinson's disease; Huntington's disease;
aplastic anaemia; ischaemia; meningitis; liver disease; Crohn's disease;
insulin-dependent diabetes; multiple sclerosis; Grave's disease; HIV;
human immunodeficiency virus; tuberculosis; lepromatous leprosy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 139..3243
FT /*tag= a
FT /note= "Human PYRIN-1 protein"
XX
PN WO200161005-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US040145.
XX
PR 17-FEB-2000; 2000US-00506067.
PR 01-SEP-2000; 2000US-00653901.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-514773/56.
DR P-PSDB; AAE07514.
XX
PT Novel polypeptide comprising NBS-1 or PYRIN-1 of the pyrin domain protein
family useful in screening and detection assays and for treating, e.g.,
cancer, viral infections, autoimmune disease, and Alzheimer's.
XX
PS Claim 1; Fig 4; 11pp; English.
XX
CC The invention relates to human NBS-1 nucleotide binding site) and PYRIN-1
protein and nucleic acids. NBS-1 and PYRIN-1 DNA, protein and their
modulators are useful in the treatment of apoptotic and inflammatory
disorders, cancer (leukaemia, melanoma, carcinoma); viral infections
(including herpesvirus and adenovirus), autoimmune diseases (systemic
lupus erythematosus (SLE), immune-mediated glomerulonephritis, arthritis)
; neurological disorders (Alzheimer's disease, Parkinson's disease,
amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, Huntington's
disease); aplastic anaemia, ischaemia, meningitis, liver diseases. NBS-1
and PYRIN-1 DNA, protein and their modulators are also used for the
treatment of inflammatory and immune disorders such as chronic
inflammatory diseases such as Crohn's disease, insulin-dependent
diabetes, organ specific autoimmunity, including multiple sclerosis,

CC Grave's disease, allergy, asthma, HIV, tuberculosis and lepromatous
CC leprosy. The present sequence is a cDNA encoding human PYRIN-1 protein
XX
SQ Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;

Query Match	25.8%	Score	800.4	DB	4	Length	3857
Best Local Similarity	58.6%	Prod. No.	4.3e-188				
Matches	1544	Conservative	0	Mismatches	1016	Indels	74
Gaps	6						
Qy	281	TGTTGAGGATCCCCAGGAAACCTACAGGGACTATGTCGCGAGGAAATTCGGGTCTATGG	340				
Db	527	TGAAGAAGATTAACGGTAAAGATACAGAAAGTACGTGAGAAACAGATTCAGTGCATG	586				
Qy	341	AAGACCGCAATGCGCGCTTAGGGAAATGTGTCAACCTCAGCCACCGGTACACCGCGTCC	400				
Db	587	AAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGCCTCAACAAACGCTTACACAGACTGC	646				
Qy	401	TGCTGGTGAAGGAGCACTCAAAACCCATGTCAGGTTCCAGCAGCAGCTTCTGGACACAGCC	460				
Db	647	GTCTCATCAGSAGCAACCGAGCACAGAGAGAGGGAGCAGGAGCTTCTGGCCATCGGC	705				
Qy	461	GGGACACGCGAGGACCGTGGGACACGAGGCTAGCCCATCAAGATAGAGACCTCTTTG	520				
Db	706	-----AAGACCAAGACGTGTGAGAGCCCCGTGAGTCCCATTAAGATGGAGTGTGTGTTG	760				
Qy	521	AGCCAGACGAGGAGCGCCCCGAGCCACCGCACCGTGTGTCATGCAAGCGCGCAGGGA	580				
Db	761	ACCCGATGATGAGCAATTTCTGAGCCTGTGCAACCGTGTGTTCCAGGGGCGCAGGGA	820				
Qy	581	TAGCAAGTCCATGCTGSCACACAAGGTGATGTGTGACTGTGGCGGACGGGAAGCTCTTTC	640				
Db	821	TTGGGAAACAACTCTGCGCCAGGAGATGATGTTGGACTGGGCATCTGGGAGCACTCTACC	880				
Qy	641	AAGCAGAGTTGATTTATCTCTTCTACATCAACTGCGAGGAGATGAACACAGAGTGCACGG	700				
Db	881	AAGCAGGTTTGACTATCTCTTCTATCTATCTGCTGTCGAGAGTGAGCC---TTGTGACAC	937				
Qy	701	AATGAGCATCAAGACCTCATTTCTAGCTGCTGGCTGTGAGCCGAGCGGCTCTTCAGG	760				
Db	938	AGAGGAGCTGGGGACCTGATCATGAGTGTGTCGCCGACCCAAACCCAGCCATCCACA	997				
Qy	761	AGCTCATCCGATTTCCGAGCGCCTCTTTTCATCATCGCGGCTTCGATGAGTCAAGC	820				
Db	998	AGATGCTGAGAAAACCTTCGAAATCTTCTCTCATGGAAGGCTTCAATGAGCTGCAG	1057				
Qy	821	CTTCTTTCCAGATCCTCAGGACCCCTGGTGCCTCTGCTGGAGGAGAAAACGGGCCACGG	880				
Db	1058	GTGCGCTTTGACGAGCACATAGACCGCTCTGACCTGACTGGCAGAAAGCCGAGCGGGAG	1117				
Qy	881	AGCTGCTCTTTAAGAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCA	940				
Db	1118	ACATTCCTCTGAGCAGCTCATCAGAAAGAGCTGCTTCCGAGGCTCTCTGCTCATCA	1177				
Qy	941	CCACAGGCCCAACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCAGGCAATGG	1000				
Db	1178	CCACGAGACCTGTGGCCCTGGAGAACTCGACACTTGTGTGACCATCTCGGCAATGG	1233				
Qy	1001	AGATCCTGGCTTCTCTGAGGACAGAAAGAGGAATATCTTCAAGATATTTCCACAATG	1060				
Db	1238	AGATCCTGGTTTCTCCGAGGCCAAAGAAAGAGTCTTCTCAAGTACTTCTCTGATG	1297				
Qy	1061	CAGAGCAGCGCGGCGAAGTCTTCAATTAAGTGGGACAAAGAGCCTCTTCTCAACATGT	1120				
Db	1298	AGGCCAAAGCCAGGCGACGCTTCAGTCTGATTCAGGAGAACGAGGTCTCTTCCACATGT	1355				
Qy	1121	GCTTCGTCCTCGTGTGTGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1180				
Db	1358	GCTTCATCCCCCTGGT	1417				
Qy	1181	GGGGGCTGTGTGAGACAGACGTCCAGGACACCACTGCAGTGTACATGTCTACCTGTGA	1240				
Db	1418	GAAGAGGCTTGCACAGACATCTAAGACACCAACCGCGTGTACGCTTCTTCTTCTTCTT	1477				

QY	1241	GTCTGATGCAACCCAAAGCCGGGGCCCGCGCCTCCAGCCCCCAACCCAAACAGAGGGGT	1300
Db	1478	GTTTGCTGCAGCCCCCGGAGGAGCCAGGACAGCGCCTCTGCGCCCACTCTGGGGGC	1537
QY	1301	TGTGCTCTTGGCGCAGATGGGCTCTGGAATCAGAAAAATCCTATTGAGGACGAGACC	1360
Db	1538	TCCTGCTCTTTGGCTGCAGATGGNACTCGNACAGAAAAATCCTGTTTGAAGGATCCGACC	1597
QY	1361	TCCGGAAGCAGCCCTPAGACGGGGGAAGACGTCTCTGCCTTCTCAACATGAACATCTTC	1420
Db	1598	TCAGGAATCATGACATGCAGAGCGCGATGTGTCTCTTCTGAGGATGAACCTGTCTCC	1657
QY	1421	AGAGGACATCAACTCTGAGAGCTACTACAGCTTCATCCACTTTGAGTTTCCAGGAATCT	1480
Db	1658	AAAGGAATGGACTCGGAGAGTCTACAGCTTCATCAGCTTCATCCACATGACTTTCCAGGATCT	1717
QY	1481	TTGAGCTATGTACTATATCTCTGAGCGAGGGGAGGGCGG-----	1520
Db	1718	TTGCCGCCATGTACTACCTGCTGGAAGAGGMAAAGNAGGAAGGACCAACGTTCACGGGA	1777
QY	1521	-----GGCAGGCCCAGACAGGAGCTCACCAGGCTGTGTACCGAGTAGCGGTTTCTG	1573
Db	1778	GTGTTTGAAGCTTCCAGCCGAGAGCTGACAGCTCTCTTGGAATCATATGGCAAAATTCG	1837
QY	1574	AAAGGAGCTTCTTGGCACTCACCAGCCGCTTCCTGTTTGGACTCTCTGAACGAGGAGACCA	1633
Db	1838	AAAGGGTATTGTATTTTGTGTGAGTTTCTCTTTGGCCTGGTAAACAGGAGAGGA	1897
QY	1634	GGAGCCACCTGGAGAGAGAGTCTCTGCTGGAGGTCTCGCGGCACTCAAGATGGACCTGT	1693
Db	1898	CCTCCTACTTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAAAATCAGGCTGGAGCTGC	1957
QY	1694	TGCAGTGGATCCAAAGCAAAGCTCAGAGCAGCGCTCCACCCTGCAGCAGGGCTCCTTGG	1753
Db	1958	TGAATGGATTTGAAGTGAAGCCAAAGCTAAAGCTGCAGATCCAGCCCAGCCAGCTGG	2017
QY	1754	AGTTCTTACGCTCTGTACAGATCCAGGAGGAGGATTTATCCAGAGGCCCTCAGCC	1813
Db	2018	AAATGTTCTACTGTTGTACGATGCAGGAGGAGGACTTCGTGCAAAAGGGCCATGGACT	2077
QY	1814	ACTTCCAGGTGATGTTGGTCAGCAACATTTGCCCTCCAGATGGAGCAGATGGTCTCCTCGT	1873
Db	2078	ATTTTCCCAAGATTGAG- --ATCAATCTCTCCACCAAGATGGACCAATGGTTTCTTCTCT	2134
QY	1874	TCTGCTGAAGCGCTGAGGAGGCCCGAGTGTGCACATTGTATGGCGCCACCTCAGCG	1933
Db	2135	TTTGCAATTGAAACTGTCATCGGCTGGAGTCACTGTCCCTGGGTTTCTCCATAACATGC	2194
QY	1934	CGGACGGGGGA-----CGCGCGAGGTGC	1959
Db	2195	CCAGGAGGAGGAGGAGGAGAAAGAGGCCGACACCTTGATATGGTGCAGTGTGTC	2254
QY	1960	TCCGACGAGCGCACACGCTGTTGGTGCAGCTCAGACCCAGAGAGGACCGTTCTGCTGGAC	2019
Db	2255	TCCCAAGCTCCTCTCATGCTGCCTGTTC-TCATGGATTGTGTGAACAGCCACCTCACTTC	2313
QY	2020	GCCTPACAGTGAACATCTGGCAGCGGCCCTGTGCAACCAATCCAACTGTATGAGTGTCT	2079
Db	2314	AGTTTTTCCGGGGCCCTCTTTTCAGTTCTCAGCACCCAGCCAGAGTCTAACTGAATGGAC	2373
QY	2080	CTGTACCGAATGCCCTGGGCAGCCGGGGGTGAAGCTGCTCTGTCAAGACTCAGACAC	2139
Db	2374	CTCAGTGACAAATCTCTGGGGGACCCAGGATGAGAGTGTGTGTGAACCGCTCCAGCAT	2433
QY	2140	CCCAACTGCAAACTTCAAGAACCTGAGGCTGAAGAGGTGCCGCATCTCCAGCTCAGCCTGC	2199
Db	2434	CCTGGCTGTAAACATTCGGAGATTGTGTTGGGGCGCTGTGGCCTCTCGCATGAGTGTCTC	2493
QY	2200	GAGGACTCTCTGCAGCTCTCATAGCCAAATGAATTTGACAGGATGGATCTCAGTGGC	2259
Db	2494	TTGACATCTCTTGTGTCCTCAGCAGCAACAGAAAGCTGTGTGGAGCTGGACCTGAGTGAC	2553
QY	2260	AACGGCGTGTGAATTTCCCAAGGCATGATGCTGCTTTGCGAGGGCCTTCGGGCAATCCCACTGC	2319

Db 2554 AACGCCCTCGGTGACTTCGGAATCAGACTCTCTGTGTGGACTGAAGCACTGTTGTGTC 2613
 Qy 2320 AGCTGCACATGATTCAGTTTGGAGAGTGTGACGTGGAGTCCGGGCTTGTGAGGATG 2379
 Db 2614 ATCTGAAGAAGCTCTGGTTGGTGTGAGTGTGCTGCTCACATCAGCATGTTGTGAGGATCTT 2673
 Qy 2380 GCTTGTGTCTCGGACCAACCCACATCTGTTGTGAGTGTGACCTGACAGGAATGCACTG 2439
 Db 2674 GCATCAGTATTGAGCACCAGCCATTCCTGACAGACTCTATGTGGGGAGATGCTTGT 2733
 Qy 2440 GAGGATTTGGGCTGAGTTTACTATGCCAGGGAGTGGAGCACCAGTCTGCGACTACGG 2499
 Db 2734 GGAGACTCAGGAGTGCATTTTATGTGAAAAAGCAAGAAATCCACAGTGTAACTTCGAG 2793
 Qy 2500 ACTTTGTGCTGAAGATCTCGGCTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2559
 Db 2794 AAACCTGGGTGTGTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2853
 Qy 2560 CTAGTGTGAACAGAGCTGAGAGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2619
 Db 2854 CTAGCACTAATCAGAACTCTAGCACTTACCTGCGAGGCAACACTCTCGAGACAAG 2913
 Qy 2620 GGGTGTCTGCTGTGTGAGGCTCAGGATCCACAGTCCAGGCTCAGGCTCAGGCTCAGG 2679
 Db 2914 GGGATCAAACTACTCTGTGAGGAGCTCTTGCACCCGAGCTGCAAGCTTCAGGTTGGAA 2973
 Qy 2680 TTGGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2739
 Db 2974 TTAGCACTGCACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3033
 Qy 2740 AACCAACTCTCGGAGCTGAGCTTGTGCTTCAAGCACTGGGAGCTGGGCTGCTGCTG 2799
 Db 3034 AGCCAGAGCTCGAAGCTGAGCTGGGCAACATGACTTGGGCACTTGGGCTGCTGCTG 3093
 Qy 2800 TTGCTGGCTGAGGCTGCAATCCGCTGAGACTCCAGAACTGTGGTGTG 2853
 Db 3094 ATGTTCTGTGAAGTGTGAACAGCAGAGCTGCTCTCTGCAAGAACTGGGCTG 3147

RESULT 10

ABX93556

ID ABX93556 standard; cDNA; 3857 BP.

XX ABX93556;

DE 22-MAY-2003 (first entry)

XX Huma cDNA encoding PYRIN-1.

Human; ss; gene; nucleotide binding site; pyrin domain; NBS-1; PYRIN-1;
 caspase-1; antiinflammatory; apoptosis; ASC; NF-kB; nuclear factor KB;
 LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria;
 arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 psoriasis; graft versus host disease; food allergy; conjunctivitis;
 chronic obstructive pulmonary disease; food allergy; conjunctivitis;
 helminthic infection; Leishmaniasis; viral infection; HIV infection;
 bacterial infection; tuberculosis; leprosy; chromosome 1q44.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 139..3243

FT /*tag= a

FT /product= "PYRIN-1"

XX US2002187922-A1.

XX PN 12-DEC-2002.

XX PD 22-APR-2002; 2002US-00127516.

XX 17-FEB-2000; 2000US-00506067.
 PR 01-SEP-2000; 2000US-00653901.
 PR 26-SEP-2001; 2001US-00964955.
 PR 20-DEC-2001; 2001US-00027629.
 XX (BERT/) BERTIN J.
 PA (MANJ/) MANJI G A.
 PI Bertin J, Manji GA;
 XX WPI; 2003-328763/31.
 DR P-PSDB; AB08503.
 XX Identification of compound that binds to polypeptide (for e.g. a PYRIN
 protein), useful for treating disorders associated with inappropriate
 apoptosis, for e.g. inflammatory disorder.
 PS Disclosure; Fig 4; 74pp; English.

XX The invention relates to a new method for the identification of a
 compound that binds to a human PYRIN-1 (an apoptotic signalling molecule
 appearing as AB08503) comprising contacting the polypeptide or a cell
 expressing the polypeptide to a test compound and determining whether
 PYRIN-1 binds to the test compound, or determining the effect of the test
 compound on the activity of PYRIN-1. In a similar manner, compounds are
 isolated which modulate the binding of PYRIN-1 to ASC (comprising testing
 the compounds against the pyrin binding domains of both PYRIN-1 and ASC
 (not defined)), modulate the activity of NF-kB (nuclear factor KB), and
 which modulate the ASC-mediated activation of NF-kB (comprising: (a)
 measuring the binding of a test compound to the LRR (leucine rich repeat)
 domain of PYRIN-1; and (b) measuring the activation of NF-kB in a cell
 expressing ASC and PYRIN-1 in the presence and absence of the LRR domain
 binding compound). The candidate modulators may be useful for treating an
 inflammatory disorder. Also included is a method (M8) for identifying a
 modulator of caspase-1 activity. The identified compound can be used to
 modulate the polypeptide's activity (ASC and NF-kB activities in a
 patient). By modulating the expression or activity of the polypeptide
 (PYRIN-1 and ASC), a disorder associated with inappropriate apoptosis
 (inflammatory disorders e.g. familial cold urticaria, arthritis,
 inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple
 sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus
 host disease, asthma, chronic obstructive pulmonary disease, allergies
 (e.g. food allergies), conjunctivitis, helminthic infection (e.g.
 Leishmaniasis), viral infections such as HIV infection, and bacterial
 infections such as tuberculosis and leprosy) can be treated. Human NBS-1
 protein (nucleotide binding site protein 1) contains a pyrin domain which
 was used to identify genomic sequences containing the PYRIN-1 gene. The
 gene for pyrin-1 is located on chromosome 1q44. The present sequence
 encodes human PYRIN-1

SQ Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;

Query Match

Best Local Similarity 25.8%; Score 800.4; DB 8; Length 3857;

Matches 1544; Conservative 0; Mismatches 1016; Indels 74; Gaps 6;

Qy 281 TGTGAGGAGTCCCGAGGAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATGG 340
 Db 527 TGAAGAAAGATTACCGTTAAGAGTACAGAAAGTACGTGAGAGCAGATTCCAGTGCATTG 586
 Qy 341 AAGACCGCAATGCGGCTTAGGGAAATGTGTCAACCTCAGGCACCGGTACACCGGCTCC 400
 Db 587 AAGACAGAAATGCGGCTCTGGGTGAGAGTGTGAGCTTCAACAAACGCTACACAGACTGC 646
 Qy 401 TGCTGGTGAAGGAGCAGCTCAACACCCATGCGAGTCCAGCAGAGCTTCTGGACACAGGCC 460
 Db 647 GTCTCATCAAGAGAGCAGCGAGCCAGCAGGAGAGGAGGAGAGCTTCTGGCCATCGGC- 705
 Qy 461 GGGGACACGCGAGGAGCAGCTGGGACACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTG 520
 Db 706 -----AAGACCAAGACGTGTGAGAGCCCGGTGAGTCCCATTAAGATGAGTGTCTTTG 760

Db 2914 GGGATCAAACTACTCTGTGAGGAGCTCTTGACCCCGACTCAAGCTTCAGGTGTGGAA 2973
 Qy 2680 TTGGCATCTGCCGCTGGCTCTCCGCTGTGAGGCTCTTTCTGTGTGCTCAGGCC 2739
 Db 2974 TTAGACAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCA 3033
 Qy 2740 AACCACAACCTCCGGAGCTGAGCTTGTAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
 Db 3034 AGCCAGAGCTTGGAAAGCTGAGCTTGGGCAACATGATCTGGGCACTTGGGGTCTATG 3093
 Qy 2800 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCGAGACTCCAGAACTGTGGCTG 2853
 Db 3094 ATGTTCTGTGAAGTCTGAAACAGCAGAGCTGCTCTCTTGCAGAACCTGGGGTTG 3147

RESULT 11

ACD27909
 ID ACD27909 standard; cDNA; 3857 BP.

XX AC ACD27909;

XX DT 23-SEP-2003 (first entry)

XX DE Human pyrin domain family protein PYRIN-1 cDNA.

XX KW Human; ss: gene; pyrin domain family; PYRIN-1; cellular differentiation;
 KW tissue typing; cellular proliferation; cell survival; apoptosis disorder;
 KW inflammatory disorder; apoptosis associated speck like protein; ASC;
 KW nuclear factor kappaB; NF-kappaB.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT CDS 139..3243
 FT /*tag= a
 FT /product= "PYRIN-1"

XX PN US2002197660-A1.

XX PD 26-DEC-2002.

XX PF 20-DEC-2001; 2001US-00027629.

XX PR 17-FEB-2000; 2000US-000506067.

XX PR 01-SEP-2000; 2000US-00053901.

XX PR 26-SEP-2001; 2001US-00064955.

XX PA (BERT/) BERTIN J.
 XX PA (MANJ/) MANJI G A.

XX PI Bertin J, Manji GA;

XX XX WPI; 2003-521543/49.

XX DR P-PSDB; AB063315.

XX PT Identifying compound that binds to polypeptide useful in treating
 PT apoptotic and inflammatory disorders, comprises contacting polypeptide
 PT with test compound and determining whether polypeptide binds to test
 PT compound.

XX PS Disclosure; Fig 4; 71pp; English.

XX CC The invention relates to a compound that binds to a polypeptide (e.g.
 CC PYRIN-1) or NBS-1. The nucleic acid molecules, proteins, protein
 CC homologues and antibodies incorporated in the invention can be used in
 CC screening assays; detection assays (e.g. chromosomal mapping, tissue
 CC typing, forensic biology); predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatments (e.g. therapeutic and prophylactic). When NBS-1 or
 CC PYRIN-1 protein interacts with other cellular proteins, they can be used
 CC for regulation of cellular proliferation and differentiation; and
 CC for regulation of cell survival. The isolated nucleic acid molecules of the

CC invention can be used to express NBS-1 or PYRIN-1 protein to detect NBS-1
 CC or PYRIN-1 mRNA or a genetic lesion in NBS-1 or PYRIN-1 gene and to
 CC modulate NBS-1 or PYRIN-1 activity. These proteins can be used to screen
 CC drugs or compounds which modulate the NBS-1 or PYRIN-1 activity or
 CC expression to treat disorders characterized by insufficient or excessive
 CC production of these proteins. The anti-NBS-1 or PYRIN-1 antibodies of the
 CC invention can be used to detect and isolate NBS-1 or PYRIN-1 proteins and
 CC modulate NBS-1 or PYRIN-1 activity. With the invention, disorders
 CC associated with inappropriate apoptosis and inflammatory disorder can be
 CC treated; and the apoptosis associated speck like protein (ASC) and
 CC nuclear factor (NF)-kappaB activity in a patient can be modulated by
 CC administering the compound that alters the activity of PYRIN-1. The
 CC method takes advantage of the interactions between NBS-1 or PYRIN-1 and
 CC various proteins involved in apoptotic and inflammatory signaling
 CC pathways. The nucleic acid molecules of the invention can be inserted
 CC into vectors. The present sequence represents cDNA encoding the human
 CC pyrin domain family protein PYRIN-1
 XX
 SQ Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;

Query Match 25.8%; Score 800.4; DB 8; Length 3857;
 Best local Similarity 58.6%; Fred. No. 4.3e-188;
 Matches 1544; Conservative 0; Mismatches 1016; Indels 74; Gaps 6;

Qy 281 TGGTGAGGGATCCCCAGGAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATGG 340
 Db 527 TGAAGAAAGATTACCGTAAGAGTACAGAAAGTACGTGAGAGCAGATTCAGTGCATGG 586
 Qy 341 AAGACCGCAATCGCGCTAGGGGAATGTCTCAACCTCAGGCACCGGTACACCCGGCTCC 400
 Db 587 AAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGGCTCAACAAACGCTACACAGACTGC 646
 Qy 401 TGCTGGTGAAGAGGACACTCAAAACCCATGCGAGTCCAGCAGAGCTTCTGGACACAGCC 460
 Db 647 GTCTCATCAAGAGGACACCGGAGCCAGCAGAGAGGAGGAGAGGAGCTTCTGGCCATCGGC- 705
 Qy 461 GGGGACACCGGAGGACCGTGGGACACACAGGCTAGGCCCATCAAGATAGAGACCTCTTTG 520
 Db 706 -----AAGACCAAGAGCTGTGAGAGCCCGCTGAGTCCCATTAAGATGAGTTGCTTTG 760
 Qy 521 AGCCAGACGAGGAGCGCCCGAGGCCACCGCGCACCGTGTCTATGCAAGCGCGGAGGGA 580
 Db 761 ACCCGGATGATGAGCATCTGAGCGCTGTGCACACCGTGTGTTCAGCGGGCGGAGGA 820
 Qy 581 TAGCGAAGTCCATGCTGGGACACAGAGTGTGAGTGGGCGGAGCGGAGAGCTTTCC 640
 Db 821 TTGGGAAAACAAATCTCTGGCCAGGAAGATGATGTTGGACTGGGACATCGGGGACACTTACC 880
 Qy 641 AAGCAGATTTGATTATCTTCTACATCAACTGCGAGGAGATGAACAGAGTGCACCG 700
 Db 881 AAGACAGGTTTGACTATCTGTTCTATATCCACTGTGCGGAGGTGAGCC---TTGTGACAC 937
 Qy 701 AATGAGATGCAAGACCTCATCTTTCAGTGTGCTGAGCCCGAGCCGCGCTCTCCAGG 760
 Db 938 AGAGGAGCCTGGGGAGCCTGATCATGAGTGTGCTGCCGACCCCAACCCACCCATCCACA 997
 Qy 761 AGCTCATCCGAGTCCCGAGCGCTCTTTCATCATGACGGCTTCGATGAGCTCAAGC 820
 Db 998 AGATCGTGAGAAAAACCCCTCCAGAAATCTCTTCTCTCATGAGCGGCTTCGATGAGTGAAG 1057
 Qy 821 CTCTTTTCCAGATCTCTCAGGAGACCTCTGCTGCTGGGAGGAGAAACCGGCCACCG 880
 Db 1058 GTGCTTTGACGAGCACATAGACCGCTCTGCACTGCTGCGAAGAGCGCGGGAG 1117
 Qy 881 AGCTGCTTCTTAACAGCTTAATTCGGAAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCA 940
 Db 1118 ACATTCTCTGAGCAGCTCATCATGAAAGAGCTGCTTCCCGAGGCTCTCTGCTCATCA 1177
 Qy 941 CCACAGGCCACCGCTTTGGAGAGCTCCACCGTCTCTGAGGAGCACCAGGAGATGG 1000
 Db 1178 CCACGAGACCTGTGGCCCTGGAGAACTGCGAGCACTTGTGACCATCTCTCGGCATGTGG 1237
 Qy 1001 AGATCTGGGCTTCTCTCAGGAGGAGAAAGGAATACTTCTTCAAGATATTTCCCAATG 1060

1238 AGATCTCTGGGTTCTCCGAGGCCAAAAGGAAGAGTACTTCTCAAGTACTTCTCTGATG 1297
1061 CAGAGCAGGCGGCAAGTCTTCAATTACGTGAGGGAACAAGAGCCTCTCTTCAACCATGT 1120
1298 AGGCCAAGCAGCAGGCGAGCTTCACTGATTCAGGAGAACAGAGTCTCTTCAACATGT 1357
1121 GCTTCGTCCTCCCTGGTGTGCTGGTGTGTGTATCTGCTCCCTCAGCAGCAGCTGAGGGTG 1180
1358 GCTTCATCCCTCGTCTGCTGGATCGTGTGCACTGGACTGAAACAGCAGATGAGAGTG 1417
1181 GGGGCTGTTCAGACAGAGCTCAGGAGCACCACCTGAGTGTATGCTCTACCTGCTGA 1240
1418 GCAAGAGCCTTGCCCAAGACATCTAAGACCAACCCGCGGTGTACGTCTTCTTCTTCCA 1477
1241 GTCTGATGCAACCAAGCCGGGGCCCGCGCTCCAGCCCCCAACCAACAGAGAGGT 1300
1478 GTTGTCTGACCCCGGGGAGGAGCAGGAGCAGGCGCTCTGCGCCACCTCTGGGGGC 1537
1301 TGTGCTCCTTGGGCGAGATGGGCTCTGGAATCAGAAATCTATTTAGAGAGCAGAAC 1360
1538 TCTGCTCTTGGTGCAGATGGAATCTGGAACCAAGAAATCCTCTTTGAGGAGTCCGACC 1597
1361 TCCGAGACAGCCTAGACGGGAAGACGCTCTGCTGCTTCTCCTCAACATGACATCTTCC 1420
1598 TCAGGAATCATGGACTCCAGAGCGGATGTGTCTGCTTTCTGAGGATGAACCTGTTC 1657
1421 AGAAGGACATCAACTGTGAGAGTACTACAGCTTCATCCACTTCAGTTTCCAGGAATCT 1480
1658 AAGAGGAGTGGACTCGAGAGATCTTACAGCTTCATCCACTTCAGTTTCCAGGATCT 1717
1481 TTGAGCTATGATATATCTCTGACGAGGGGAGGCGG----- 1520
1718 TTGCGCCATGTACTACTCTGTGAAGAGGAAAGGAAGAACGATCCAGGA 1777
1521 -----GGCAGGCCAGACAGACGCTGACGAGCTGTGACCGAGTACGCTTTCTG 1573
1778 GTGCTTTGAGCTTCCAGCGGAGAGCTGACAGTCTTCTGGAATCTATGGCAATTCG 1837
1574 AAAGGAGCTTCTGGCACTCACCAGCGCTTCTCTTTGGAATCTCCTCAACGAGGAGACCA 1633
1838 AAAGGGTATTTGATTTTGTGATCGTTTCTCTTTGGCTGTAAACAGAGAGGA 1897
1634 GAGGCCACTGGAGAGAGTCTCTGTGGAAGTCTCGCGCACATCAAGATGACCTGT 1693
1898 CCTCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGAAATCAGGCTGAGCTGC 1957
1694 TGAGTGGATCCAAAGCAAGCTCAGAGCGAGCTCCACCTGACAGGGCTCCTTGG 1753
1958 TGAATGGATTGAAGTGAAGCCAAAGCTAAAGCTGCAAGATCCAGCCAGCGAGCTGG 2017
1754 AGTTCTTCAGCTGCTGTACAGATCCAGAGGAGGAGTTTATCCAGCAGGCGCTGAGCC 1813
2018 AATTGTTCTACTGTTGTACAGATGACAGAGGAGGACTTCGTGCAAGGGCCATGGACT 2077
1814 ACTTCAGGTGATGCTGGTCAAGCAATTCGCTCCAGATGGAGCAATGCTCTCTCGT 1873
2078 AATTCCCAAGATTGAG---ATCAATCTCTCCACAGATGGACCAATGTTTCTTCTCT 2134
1874 TCTGCTGAGAGCGCTGAGGAGCGCCAGGCTGCTGACCTGTATGGCGCCACCTACAGG 1933
2135 TTTGCAATTGAACATGTCTCGGTGAGTCACTGTCTCTGGGTTTCTCCATTAACATGC 2194
1934 CGACCGGGGAAGA-----CCGCGCGAGGTGC 1959
2195 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2254
1960 TCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2019
2255 TCCCAAGCTCTCTCATGCTGCTGCTTTC-TCATGGATTGGTGAACAGGCCACCTCACCTCC 2313
2020 GCCTACAGTGAACATCTGGCAGGCGGCTGTGACCAATCCAAACCTGATAGACTCTCT 2079

2314 AGTTTTCCCGGGGCTCTTTTTCAGTTCTGAGCACCAGCAGAGTCTAACTGAATTGGAC 2373
2080 CTGTACCAAATGCCCTGGGCGAGCGGGGTGAAGCTGTCTCTCAAGACTCAGACAC 2139
2374 CTAGTGCACATTTCTTGGGGGACCCAGGAGTGAAGTGTGTGAAACGCTCCAGCAT 2433
2140 CCCAACTGCAAAACCTCAGAACCTGAGGCTGAAGAGGTCCGCAATCTCCAGCTCAGCCTGC 2199
2434 CTTGGCTGTAACATTCGAGATTTGGTGGGGCGCTGTGGCCCTCTCGCATGAGTGTGTC 2493
2200 GAGGACCTCTCTGCAGCTCTCATAGCCAAATAAGATTTGACAAGATGGATCTCAGTGGC 2259
2494 TTCACATCTCTTGGTCTCAGCAGCAACAGAAAGCTGTGGAGCTGGACCTCAGTGCAC 2553
2260 AACGGCTTGGATTTCCAGGCAATGATGCTGCTTTGCGAGGGCCTGCGGATCCCCAGTGC 2319
2554 AACGCCCTCGGTGACTTCGGAAATCAGACTTCTGTGTGGGACTGAAGCACTTGTGTGC 2613
2320 AGGCTGCAGATGATTCAGTTGAGGAAGTGTAGCTGGAGTCCGGGGTGTGTGAGAGATG 2379
2614 AATCTGAAGAAGCTCTGTGTGGTCAAGCTGTGCTCATCATCAGCATGTTGTGAGGATCT 2673
2380 GCTTCTCTGCTCGCACCAACCAATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2439
2674 GCATCAGTATTGAGCACCAGCCATTCCTCTGACCAAGACTCTATGTGGGGGAAATGCTTTG 2733
2440 GAGGATTTGGCCCTGAGGTTACTATGCGAGGACTGAGGACCCAGCTGTGAGACTACGG 2499
2734 GGAGACTCAGAGTTCGCAATTTTATGTGAAAGCCAGAAATCCACAGTGTAACTGCAG 2793
2500 ACTTGTGGCTGAAGATCTGCCGCTCACTGCTGCTGCTGTGAGAGCTGGCCTCAACT 2559
2794 AAAGTGGGGTGTGTGAATTTCTGGCTTACGCTCAGTCTGTGTGTTGAGCTTGTCTCTCGTA 2853
2560 CTGAGTGTGAACAGAGCTCAGAGAGCTGAGAGCTGAGCTGAGCTGAATGAGCTGGGAGCTC 2619
2854 CTGAGCACTAATCAGAAATCTCAGCACTTTACCTGCGAGGCAACACTCTCGGAGAACG 2913
2620 GGGTGTGCTGTGTGTGAGGGGCTCAGGCTCAGGCTCCCAAGCTCCAGACCTGCGG 2679
2914 GGGATCAAACTACTCTGTGAGGACTCTTGACCCCGACTGCAAGCTTCAGTGTGGAA 2973
2680 TTGGGCAATCTGCCGGCTGCGGCTGTGAGGGTCTTCTGTGTGAGTCTCCAGGCC 2739
2974 TTAGCAACTGCAACCTCAGCTCAGCTGTGCTGGGATCTTTCCACACTTCTGACCTCC 3033
2740 AACCAACACTCCGGGAGCTGGACTTGAAGTTTCAACACCTGAGAGCTGGGCTGTGG 2799
3034 AGCCAGAGCTTGGAAAGCTGAGCTGGGCAACAAATGACCTGGGCGACCTGGGGGTGATG 3093
2800 TTGTGCTGAGGGGCTGCAACATCCCGCTGCGAGCTCCAGAACTGTGCTG 2853
3094 ATGTTCTGTGAGTGTGTGAACAGCAGAGCTGCTCTCTGCAAGACCTGGGGTGTG 3147

RESULT 12
ADE36451
ID ADE36451 standard; cDNA; 2524 BP.
XX ADE36451;
AC ADE36451;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human PAAD domain associated protein cDNA.
XX
KW cytosolic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
KW antiallergic; antitumor; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pvrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;

cytokine receptor signaling caspase-mediated proteolysis;
 c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
 inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
 neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
 fibrosis; smooth muscle cell proliferation; balloon angioplasty;
 restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
 arthritis; lupus; schrojen's syndrome; Crohn's disease;
 ulcerative colitis; graft versus host disease; stroke; heart failure;
 neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
 cancer therapy; PAAD domain family; human; pyrin 2; gene; ss;
 PAAD domain.

Homo sapiens.

US2003077699-A1.

24-APR-2003.

25-SEP-2001; 2001US-00965621.

26-SEP-2000; 2000US-00671760.

26-SEP-2000; 2000US-0367367P.

(REED/) REED J C.

(GODZ/) GODZIK A.

(CHUZ/) CHU Z.

(PAWL/) PAWLOWSKI K.

(FIOR/) FIORENTINO L.

(ARIZ/) ARIZA M E.

(STEH/) STEHLIK C.

Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME,
 Stehlik C;

WPI: 2002-471256/50.

P-PSDB; ADE36452.

Novel isolated PAAD domain containing polypeptide useful for inducing
 apoptosis by inhibiting nuclear factor kappa B activation and in gene
 therapy for treating cancer.

Disclosure; SEQ ID NO 58; 93pp; English.

The invention describes an isolated PAAD domain containing polypeptide
 (I) comprising 80% identity to the amino acid sequence of PAAD and
 nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 -like protein containing a caspase recruitment domain (ASC)-2 fully
 defined in specification, where (I) is biologically active. (I) is useful
 for identifying a (I)-associated polypeptide, an agent altering that
 association and agents that modulate PAAD domain mediated inhibition of
 nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
 for identifying an agent that modulates the activity of the NB-ARC domain
 of (I). (I) or its functional fragments is useful in altering cellular or
 biochemical process such as apoptosis, NFkappaB induction, cytokine
 processing, cytokine receptor signaling caspase-mediated proteolysis or c
 -Jun N-terminal kinase activation, thus having modulating effect on cell
 life and death (apoptosis) inflammation, cell adhesion or other cellular
 or biochemical processes. (I) is useful for treating cancer pathologies,
 keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 proliferation in arteries following balloon angioplasty (restenosis),
 leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 versus host disease, stroke, heart failure, neurodegenerative diseases
 such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 therapy. This sequence encodes a human PAAD domain associated protein.

Sequence 2524 BP; 590 A; 580 C; 637 G; 564 T; 0 U; 153 Other;

Query Match

Best Local Similarity 20.6%; Score 639; DB 6; Length 2524;

Matches 1069; Conservative 0; Mismatches 585; Indels 39; Gaps 4;

KW	QY	281	TG	GTGAGGATCC	CGAGAACT	CTACAGGACT	ATGTCCCGAGGAAT	TCGGCTCATG	340
KW	DB	570	TG	AGAAAGATT	ACCGTA	AGAGTAC	AGAAAGTAC	GTGAGAGCA	629
KW	QY	341	AA	GACCGCAAT	CGCGCCT	TAGGGGAAT	GTGTCAACCT	CAGCCACCGGT	400
KW	DB	630	AA	GACGGAAT	CGCGTCT	TGGGTGAG	GTGTGAGCTT	CAACAACT	689
KW	QY	401	TG	TGTTGAAG	GAGCACT	CAAAACCC	CATGCAGG	TCCTTCTTGG	460
KW	DB	690	GT	CTCATCA	AGGAGCA	CCCGAGC	AGGAGGAG	GAGGAGGCT	748
OS	QY	461	GG	GGACACG	CGGAGG	ACCGGT	GAGCCAT	CAAGTAGAG	520
PN	DB	749	-----	AA	GACCAAG	AGTGTG	AGAGCC	CGGTAGT	803
PD	QY	521	AG	CCAGAG	GAGGCG	CCCGGAG	CCACCG	CGTGTG	580
PF	DB	804	AC	CCCGAT	GATGAG	CAATCT	GAGCC	TGTGCA	863
PR	QY	581	TA	GGCAAG	TCTCAT	GTGGCA	CAAAAG	TGTGAG	640
PT	DB	864	TT	GGAAAA	CAATCT	CTGGCC	AGGAAG	ATGATG	923
PA	QY	641	AA	GGCAGAT	TTTGA	TATCTCT	TCTACAT	CACTG	700
PA	DB	924	AA	GACAGT	TTGACT	ATCTG	TCTAT	CTCTG	980
PI	QY	701	AA	TGAGAT	GAAGCA	CTCATCT	TTCAGT	GTGGCT	760
PI	DB	981	AG	AGAGG	CCCTGG	GACCTG	ATCATG	AGCTG	1040
DR	QY	761	AG	CTCAT	CGAGT	TTCCG	AGGCGC	CTCTT	820
DR	DB	1041	AG	ATCGT	GAGAAA	ACCCTC	CAAGAT	CTCTT	1100
PT	QY	821	CT	TCCTT	CCAGCAT	CCTCAG	GGACCT	GTGTG	880
PT	DB	1101	GT	GCCTT	GGACGAC	ATAGG	ACCGCT	CTGCA	1160
PS	QY	881	AG	CTGCTT	CTTAA	CAGCTT	TAATTC	GGAAG	940
CC	DB	1161	AC	ATTCTC	CTGTAG	CAGCTC	ATCAGA	AGAGT	1220
CC	QY	941	CA	CACG	CCGCA	CGCTTT	GGAGAG	CTCC	1000
CC	DB	1221	CC	ACGAG	ACCTG	TGGCC	CTTGG	AGAACT	1280
CC	QY	1001	AG	ATCTT	CGGCTT	CTCTG	CAGGCA	AAAGGA	1060
CC	DB	1281	AG	ATCTT	GGGTTT	CTCCG	AGGCA	AAAGGA	1340
CC	QY	1061	CAG	AGC	AGGCGG	CCCAAG	CTTTC	CAATAC	1120
CC	DB	1341	AG	GC	CAAGC	AGCAG	CGCTT	CACTT	1400
CC	QY	1121	GT	TGCT	CCGCTT	GTGTG	TGTG	TGTA	1180
CC	DB	1401	GT	TCAT	CCCC	CTGTG	TGCTG	GAATG	1460
CC	QY	1181	GG	GGGCTG	TTGAG	CAGAG	CGTCC	AGGAC	1240
CC	DB	1461	GA	AGAG	CTTGG	CCGAG	ACATCC	AGAG	1520
CC	QY	1241	GT	CTGAT	CAAC	CCCAAG	CGGGG	CCCC	1300
CC	DB	1521	GT	TTGCT	GAC	CGCCG	GGGAG	CGGAG	1580
CC	QY	1301	TG	TGCT	CTCTG	GCGC	AGAT	GGCT	1360
CC	DB	1581	TCT	GCTCT	TTGG	CTGC	AGAT	GAAT	1640

DR P-PSDB; ABR01160.
XX
PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
PS
XX Claim 21; Page 1274-1275; 2474pp; English.
CC
CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cell differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification, chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein-encoding cDNA clone of the invention
XX
SQ Sequence 2847 BP; 706 A; 681 C; 719 G; 740 T; 0 U; 1 Other;

Query Match 19.1%; Score 594.6; DB 7; Length 2847;
Best Local Similarity 57.1%; Pred. No. 4.9e-137;
Matches 1203; Conservative 0; Mismatches 839; Indels 65; Gaps 4;

QY 808 GATGAGCTCAGGCTCTTCCAGATCTCTCAGGACCTCAGGACCTCTGCTCTGGAGGAG 867
DB 8 GATGAGCTCAGGCTCTTCCAGATCTCTCAGGACCTCAGGACCTCTGCTCTGGAGGAG 67

QY 868 AAACGGCCACGAGGAGTGTCTTAAAGCTTAACTCGGAGAGAGCTGCTCCCTGAGCTA 927
DB 68 GCCGAGGGGGAGACATCTCTGAGCAGCTCATCAGAAAGAGCTGCTTCCGAGGCC 127

QY 928 TCTTTGCTCATACACAGCGCCAGCTTTGAGAGGCTCCACCGTCTGCTGGAGCAC 987
DB 128 TCTTCTCTCATACACAGCGACCTGTGGCCCTGGAGAACTGCGACACTTGTCTGGACAT 187

QY 988 CCCAGGATGTGAGATCTCTGGGCTTCTGAGGAGAGAGAGAGGATATCTTCAAG 1047
DB 188 CTTGGCATGTGAGATCTCTGGGCTTCTGAGGAGAGAGAGAGAGATATCTTCAAG 247

QY 1048 TATTTCCACAATCAGAGAGCGGGCCAAAGTCTTCAATACGTGAGGGACAACAGGCT 1107
DB 248 TACTTCTGATGAGGCCCAAGCAGGCGAGCTTCACTGATTCAGGAGACGAGGTC 307

QY 1108 CTCTTCAACATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167
DB 308 CTCTTCAACATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367

QY 1168 CAGCTGGAGGCTGGGGCTGTTGAGACAGACCTTCCAGGACCACTGAGGTGATCATG 1227
DB 368 CAGATGGAGGTGGAGAGGCTTGGCCAGACATCCAAAGACCACTGAGGTGATCATG 427

QY 1228 CTCTACTGTGATCTGATGACCAACCAAGCCGCGGGGCCCGCGCTTCCAGCCGCCACCC 1287
DB 428 TTCTTCTTCCAGTGTGCTGAGCCCGCGGGAGGGAGCCAGGACAGCGCTCTGGCC 487

QY 1288 AACCAAGAGGCTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAAATCTTATTT 1347
DB 488 CACCTCTGGGGGCTCTGCTCTTGGGCTGAGATGGAATCTGGAACCAAGAAATCTCTGTT 547

QY 1348 GAGGAGCAGGACCTCCGGAAGCAGCGCTTACAGCGGGAGAGAGCTCTCTGCTTCTCAAC 1407

DB 548 GAGGAGTCCGACCTCAGGAATCATGGACTGCAGAAGCGGATGTGTCTGCTTCTCGAGG 607
QY 1408 ATGAACATCTTCCAGAGGACATCACTGTGTGAGAGGTACTACAGCTTTCATCCACTTCAGT 1467
DB 608 ATGAACCTGTTCMAAAGGAAGTGGACTTGGAGAGTTCACAGCTTTCATCCACATGACT 667
QY 1468 TTCAGGAATCTTTGTCAGCTATGTACTATATCTCTGAGCAGGGGGAGGGCGG----- 1520
DB 668 TTCAGGAGTCTTTGCGCCATGTACTACTCTGTGGAAGGAAAGGAAGGAGGACG 727
QY 1521 -----GGCAGGCCACAGACCGAGCTTCCAGCGGACGTCGACAGTCTTCTGGAAGAC 787
DB 728 AACGTTCCAGGAGTCTGTTGAAGCTTCCAGCGGACGTCGACAGTCTTCTGGAAGAC 787
QY 1561 TAGCGGTTTCTGAAAGGAGCTCTCTGCACTCACCAGCGCTTCCTGTTTGGACTCTCTG 1620
DB 788 TATGGCAATTCMAAAGGGGTATTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 847
QY 1621 AACGAGGAGACCGAGGACCTCTGGAAGAGTCTCTGCTGGAAGGTCTCCGCGCACATC 1680
DB 848 AACGAGGAGGAGCTCTCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAAAATC 907
QY 1681 AAGATGAGCTGTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTCGAC 1740
DB 908 AGGCTGAGCTGTGAAATGGATTGAAGTGAAGCCAAAGCTAAAGAGCTCGAGATCCAG 967
QY 1741 CAGGGCTCTCTGAGTCTTTCAGCTGCTCTGAGATCCAGGAGGAGGAGTTTATCCAG 1800
DB 968 CCAGGAGCTGGAATGTTCTACTGTTGATGATGATGATGATGATGATGATGATGATGATG 1027
QY 1801 CAGGCTCTGAGCCACTTCCAGGTGATGCTGCTGAGCAACATTCCTCCAAAGATGAGGAC 1860
DB 1028 AGGCGCATGAGCTATTTCCCAAGATTGAG--ATCAATCTCTCCACAGAGTGGACAC 1084
QY 1861 ATGCTCTCTGCTCTGCTGTAAGCGCTGAGAGCGCCAGGCTGCTGCTGCTGCTGCTGCTG 1920
DB 1085 ATGCTTCT 1144
QY 1921 GCCACCTTACAGCGGAGCGGGAAGA----- 1946
DB 1145 CTCTATAACATGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1204
QY 1947 CCGCGGAGGTGTCTCGAGGAGCGACAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2006
DB 1205 GTGCAAGTGTCTCTCCAAAGCTCTCTCATGCTGCTGCTCTCAT--GGAATTGTTGAACAG 1263
QY 2007 CGTCTCTGAGCGCTTACAGTGAACATCTGCGAGCGGCTCTGCTGCTGCTGCTGCTGCTG 2066
DB 1264 CCACCTTCACTTCCAGTCTTTTGGCGGGGCTCTTTTCACTTCTGAGCACCGAGGAGTCT 1323
QY 2067 GATAGAGCTGTCTCTGTACCGAAATGCCCTGGGCGAGCGGGGGGTGAAGCTGCTCTGCTCA 2126
DB 1324 AACTGAATGGAGCTCAGTGACAAATCTCTGGGGGAGCCAGGAGTGAAGTGTGTGTGA 1383
QY 2127 AGGACTCAGACACCCCACTGCAAACTTCAGAACCTCAGAGCTGAAGAGTGCCTCATCTC 2186
DB 1384 AACGCTCCAGCATCTCTGGGTGTAACATTCGAGATTTGTTGGGGCGCTGTGGCTCTC 1443
QY 2187 CAGCTCAGCTCGAGGAGCTCTCTGAGCTCTCATAGCCAAATGAAGATTTGACAGGAT 2246
DB 1444 GCATGAGTGTCTTTCGACATCTCTTGGTCTCTGAGCAGCAACAGAGCTGTGTGAGCT 1503
QY 2247 GATCTCAGTGGCAACGGGTTTGGATTCCAGGATGATGCTCTTTCAGGAGGCTGCG 2306
DB 1504 GGACCTGAGTGACAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563
QY 2307 GCATCCCGAGTGCAGGCTCAGATGATTCAGTTGAGGAAGTGTGCTGAGAGTGCCTGAGG 2366
DB 1564 GCACCTGTTGTGAATCTTGAAGAGCTCTGTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1623
QY 2367 TTCTCAGGAGATGGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2426

Db 1624 TTGTCAGGATCTTGCAATCAGTATTGAGCACCAGCCATTCCTCCAGCAGCTCTATGTGGG 1683
QY 2427 AGGAATCACTGGAGGATTTGGGCTCAGGTTACTATGCCAGGAGCTGAGGCCACCACT 2486
Db 1684 GGAGATGCTTGGAGACTCAGGAGTCGCAATTTTATGTGAAAAGCAAGATCCACA 1743
QY 2487 CTGCAGACTAGGACTTTGTGGCTGAAGATCTGCCGCTCACTGTGCTGTGCTGAGCA 2546
Db 1744 GTGTAACTGCAGAACTGGGTTGGTGAATCTGGCCCTTACGTCAGTCTGTTGTCAGC 1803
QY 2547 GCTGGCTCACTCTCAGTGTGAACAGAGCTGAGAGCTGGAGCTGGAGCTGAATGA 2606
Db 1804 TTTGTCCTCGTACTCAGCACTAATCAGAACTCTCAGCAGCTTTACCTGCGAGGCAAC 1863
QY 2607 GCTGGGACCTCGGGTGTCTGCTGTGTGAGGCTCAGGCATCCACGTCGAAGCT 2666
Db 1864 TCTCGAGACAGGGGATCAAACTACTCTGTGAGGAGCTTTGACACCCGCTGCAAGCT 1923
QY 2667 CCAGACCTCGGGTTGGGCTCTGCCGCTGGGCTCTGCCGCTGTGAGGCTCTTTCTGT 2726
Db 1924 TCAGGTGTGGAATTAGACAACCTCACTCAGCTCAGCTGCTGTGGATCTTTCCAC 1983
QY 2727 GGTGCTCAGGCCCAACCACTCCGGAGCTGAGCTGAGCTTTCAAGCAGCTGGGAGA 2786
Db 1984 ACTTCTGACCTCCAGCAGAGCTCGAAAGCTGAGCCTGGGCAACAATGACCTGGGGA 2043
QY 2787 CTGGGCTGTGTTGCTGGCTGAGGGCTGCAACATCCGCTGAGAGCTCCAGAACT 2846
Db 2044 CTTGGGCTCATGATGTTCTGTGAAGTCTGAAACAGCAGAGCTCCCTGCAAGCT 2103
QY 2847 GTGGCTG 2853
Db 2104 GGGGTG 2110

RESULT 15
ADA98038
ID ADA98038 standard; cDNA; 2847 BP.
XX
XX
AC ADA98038;
XX
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein cDNA sequence #132.
KW human; secreted protein; cardiovascular disorder; arrhythmia;
KW atherosclerosis; stroke; endocarditis; congestive heart failure;
KW rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;
KW migraine; thrombosis; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; renal disorder; proliferative disorder; cancer; gene;
XX ss.
XX
OS Homo sapiens.
XX
XX
FN WO2003004623-A2.
XX
XX
PD 16-JAN-2003.
XX
XX
PF 26-MAR-2002; 2002WO-US009922.
XX
XX
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-247946/24.
XX
XX
PT New human secreted polypeptide and nucleic acid molecules, useful for
diagnosing, preventing, prognosticating or treating cardiovascular

disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or thrombosis).
Claim 1; SEQ ID NO 142; 1572pp; English.
The invention comprises the amino acid and coding sequence of human secreted proteins. The DNA and protein sequences of the invention are useful in the treatment of cardiovascular disorders, such as: arrhythmia, atherosclerosis, stroke, endocarditis, congestive heart failure, rheumatic heart disease, cardiomyopathy, haemorrhoids, varicose veins, migraine, or thrombosis. The DNA and protein sequences may also be used for treating or preventing: neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, renal disorders, proliferative disorders and/or cancerous diseases. The present cDNA sequence encodes a human secreted protein of the invention. NOTE: The present sequence is shown on the WIPO website.
Sequence 2847 BP; 706 A; 681 C; 719 G; 740 T; 0 U; 1 Other;
Query Match 19.1%; Score 594.6; DB 7; Length 2847;
Best Local Similarity 57.1%; Pred. No. 4.9e-137;
Matches 1203; Conservative 0; Mismatches 839; Indels 65; Gaps 4;
QY 808 GATGAGCTCAAGCCCTTCTTCCAGCATCTCTCAGGACCTCTGCTCTGCTGGAGGAG 867
Db 8 GATGAGCTGCAAGTGCCTTTGACGAGCACATAGGACCGCTCTGCACCTGACTGGCAGAAG 67
QY 868 AAACGGCCACGAGCTGCTCTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTA 927
Db 68 GCCAGCGGGGAGACATTTCTCTGAGCAGCTCATCAGAAAGAGCTGCTTCCCGAGGCC 127
QY 928 TCTTTGCTCATCACACACGCGCCCTTTGGAGAGCTCCACCGCTCTGCTGGAGCAC 987
Db 128 TCTTGTCTCATCACACGAGACCTGTGGCCCTTGGAGAACTGCACACTTGTCTGGACCAT 187
QY 988 CCCAGGCATGTGAGATCCTGGGCTTCTTCTGAGGAGAAAGAGAACTACTTCTACAG 1047
Db 188 CTTGGCATGTGGAGATCTTGGTTTCTCCAGGCGCAAGAGAAAGAGTACTTCTTCAAG 247
QY 1048 TATTTCACATGACAGCAGCGGGCCCAAGTCTTCAATTAGTGAGGAGCAACAGAGCT 1107
Db 248 TACTTCTCTGATAGGCGCCCAAGCGGCGAGCGCTTCACTCTGATTCAGAGAAACGAGGTC 307
QY 1108 CTCTTCAACATGTCTTCTCCCTGCTGCTGGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1167
Db 308 CTCTTCAACATGTCTTCTCCCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 367
QY 1168 CAGCTGGAGGTGGGGGCTGTGTGAGACAGAGCTCCAGGACCCACCACTGCACTGTACATG 1227
Db 368 CAGATGGAGGTGGCAAGAGCTTGTCCAGACATCCAGACCCACCCCGGTGTACGTC 427
QY 1228 CTCTACCTGTGAGTCTGATGCAACCCAGCGGGGGCCCGCGCTCTCAGGCCCCACCC 1287
Db 428 TTTCTTCTTTCAGTTTGTCTGAGCGCGGGAGGAGCGAGCAGCGGCTCTGCGCC 487
QY 1288 AACAGAGAGGTTGTGTCTTCTTGGCGCAGATGGGCTCTGGAATCAGAAATCTATT 1347
Db 488 CACTCTGGGGCTCTGCTCTTTGGCTGAGATGGAATCTGGAACAGAAATCTCTGTT 547
QY 1348 GAGGAGCAGGACCTCCGGAAGCAGCGGCTAGACGGGGAAGAGCTCTGCTCTCTCTCAAC 1407
Db 548 GAGGAGTCCGACCTCAGGAATCATGAGCTGAGAGGCGGATGTGTCTGCTTCTGAGG 607
QY 1408 ATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCACTTCACTTGA 1467
Db 608 ATGAACCTGTTCAGAAAGAGTGTGACTGCGAGAGTTCTACAGCTTCACTCACTGACT 667
QY 1468 TTCCAGGAACTTCTTGGCAGCTGTACTATATCTCTGAGCGGGGGGAGGCGG----- 1520
Db 668 TTCCAGGAGTCTTTTGGCGCCATGACTACTCTGCTGGAGAGAGAAAGAGAGGAGCG 727
QY 1521 -----GGCAGGCCAGACCCAGGAGCTGACCGAGCTGTGACCGAG 1560

Db 728 AACGTTCCAGGGAGTCTTTGAAGCTTCCAGCGACGTGACAGTCTTCTGGAAGAAC 787
Qy 1561 TACCGCTTTTCTGAAGAGCTTCTCGGCACTACACAGCGCTTCTCTGTTTGGACTCTGT 1620
Db 788 TATGGCAATTCGAAAGGGTATTTGATTTTGTGTTAGTCTTCTCTTGGCTGGTA 847
Qy 1621 AACGAGAGACGAGGACCTGAGAGAGTCTCTGTGGAAGTCTCGCCGACATC 1680
Db 848 AACGAGAGAGGACCTCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATC 907
Qy 1681 AAGATGACCTGTGTGAGTGGATCCAAAGCAAGCTCAGAGCGAGCTCCACCTGCAG 1740
Db 908 AGGCTGAGCTGTCTGAAATGGAATGGAAGTGAAGCCAAAGCTAAAGAGCTGCAGATCCAG 967
Qy 1741 CAGGGCTCTTGGAGTCTTTCAGCTGTGTACAGATCCAGAGGAGGAGTCTTATCCAG 1800
Db 968 CCCAGCCAGCTGGAATTTCTACTGTGTGTACAGATCCAGAGGAGGACTTCTGTCAA 1027
Qy 1801 CAGCCCTGAGCCACTTCCAGGTGATCGTGTGAGCAACATTTGCTCCAGATGGAGAC 1860
Db 1028 AGGCCCCAGTACTATTTCCCAAGATTGAG---ATCAATCTCTCCACAGAAATGGACCAC 1084
Qy 1861 ATGCTCTCTCTGTTCTGTCTGAGCGCTGCAGGAGCGCCAGGCTGCTGCACTGTATGGC 1920
Db 1085 ATGTTTCTTCTTTCATGAGAACTGTATCGGGTGGAGTCACTGTCCCTGGGGTTT 1144
Qy 1921 GCCACCTACAGCGCGGACGGGAAGA----- 1946
Db 1145 CTCATTAACATGCCAAGGAGGAGAGAGGAGGAGAAAGAGCCGACCTTGATATG 1204
Qy 1947 CCGGGGAGTGTCTCCAGAGAGGAGGACACGCTGTTGTCAGCTCAGACAGAGAGGAC 2006
Db 1205 GTGCGAGTGTCTCTCCCAAGCTCTCTCATGCTGCTGTTTCAT-GGATTTGTTGAACAG 1263
Qy 2007 CGTTCTCTGAGCCCTACAGTGAACATCTGGCAGCGCCCTTGACCAATCCAAACCT 2066
Db 1264 CCACCTCACTTCAGTTTTTGGCGGGCCCTCTTTTCAGTTCTGAGCACCAGCCAGAGTCT 1323
Qy 2067 GATAGAGCTGTCTGTGACGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGTCTGTCA 2126
Db 1324 AACTGAATTGACCTCAGTGACAAATCTCTGGGGGACCCAGGGATGAGAGTGTGTGA 1383
Qy 2127 AGACTCAGACACCCCACTGCAAACTTCAGAACTGAGGCTGAAGGTCGCGCATCTC 2186
Db 1384 AACGCTCCAGCATCTCGCTGTAACTTCGAGATTTGTTGGGGCGCTGTGCGCTCTC 1443
Qy 2187 CAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCCAAATAAGAAATTTGACAGGAT 2246
Db 1444 GCATGAGTGTGTTGACATCTCTTGGTCTCAGCAGAAACAGAACTGTGTGAGCT 1503
Qy 2247 GGATCTCAGTGGCAACGGCTTGGATTCAGAGCATGATGCTGTTTGGAGGGCTTGGC 2306
Db 1504 GGACCTGAGTGACAAAGCCCTCGTGTACTTCGGAATCAGACTCTGTGTGGGACTGAA 1563
Qy 2307 GCATCCCGAGTGAAGCTCAGATGATTCAGTTGAGGAATGTGAGTGGAGTCCGGGGC 2366
Db 1564 GCACCTGTTGTGAAATCTGAAGAGCTCTGTTGGTGTGAGTGTGCTCATCAGCATG 1623
Qy 2367 TTCTCAGGAGATGGCTTCTGTGCTCGGACCAACCCACATCTGTTGAGTTGGACCTGAC 2426
Db 1624 TTGTGAGATCTTGCATCAGTATTGAGCACAGGCAATCCCTGACAGACTCTATGTGGG 1683
Qy 2427 AGAAATGCACTGAGGAGATTGGCCCTGAGGTTACTATGCCAGGACTGAGGCCACCCAGT 2486
Db 1684 GGAGAAATGCCCTTGGGAGACTCAGAGTCCGAAATTTATGTGAAAGCCAAAGATCCACA 1743
Qy 2487 CTGCACTACGAGACTTTGTGGCTGAGATCTGCCGCTCAGTGTCTGCTGTGAGCA 2546
Db 1744 GTGTAACCTGCAGAAACTGGGGTGTGGAATTTCTGGCTTACGCTCAGTCTGTGTTCCAGC 1803
Qy 2547 GCTGGCCTCAACTCTCAGTGTGAACACAGAGCTGAGAGAGCTGGACCTGAGCCTGAATGA 2606

Db 1804 TTTGTCTCTCGGTACTCAGCACTAATCAGAATCTCACGCACTTTTACTGTGAGGCAACAC 1863
Qy 2607 GCTGGGGGACTCTGGGGTGTCTGTGTGTGAGGGCTCAGGCATCCCACTGCAAGCT 2666
Db 1864 TCTCGGAGACAGGGGATCAAACTACTCTGTGAGGACTCTTGACCCCGACTGCAAGCT 1923
Qy 2667 CCAGACCTCGGGTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGT 2726
Db 1924 TCAGGTGTTGGAATTAGACAACCTGCAACCTCACGTCACTGTCTGTGGGATCTTTCCAC 1983
Qy 2727 GGTCTCCAGGCCAACCAACCTCCGGGAGCTGGAATTTGAGTTTCAACGACCTGGGAGA 2786
Db 1984 ACTTCTGACCTCCAGCCAGAGCTTCCGAAAGCTGAGCTGGGCAACAAATGACCTGGGCGA 2043
Qy 2787 CTGGGGCTCTGTGTTGTGTGAGGGCTGCAACATCCCGCTTGCAGACTCCAGAAACT 2846
Db 2044 CCTGGGGTCAATGATGTTCTGTGAGTGTGAAACAGCAGAGCTGCTCTCTGCAAGACCT 2103
Qy 2847 GTGGCTG 2853
Db 2104 GGGGTTG 2110

Search completed: August 8, 2004, 01:41:11
Job time : 1147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 07:08:21 ; Search time 11827 Seconds
(without alignments)
11390.052 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctagcaacgcaggcag.....attggacattggctgctga 3108

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.rod.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3108	100.0	3108	6	AX417214 Sequence
2	3057	98.4	3300	6	AX459869 Sequence
3	2820	90.7	3731	9	AY116204 Homo sapi
4	2577	82.9	3507	9	BC028069 Homo sapi
5	2562	82.4	3563	9	AY116205 Homo sapi
6	2475	79.6	3827	9	AY095146 Homo sapi
7	2475	79.6	3827	9	AY154467 Homo sapi
8	2223	71.5	3395	9	AY116206 Homo sapi
9	2220	71.4	3221	9	AY116207 Homo sapi
10	1995	64.2	3466	6	AX833594 Sequence
11	1995	64.2	3466	9	AC095460 Homo sapi
12	1703	54.8	147330	9	AC008753 Homo sapi
13	1356	43.6	2158	6	AX575503 Sequence
14	768	24.7	4931	6	AX684291 Sequence
15	515	16.6	2494	9	AF231021 Homo sapi
16	143	4.6	308	6	AX684295 Sequence
17	69	2.2	80	6	AX684299 Sequence
18	28	0.9	2159	10	AY364010 Mus muscu
19	28	0.9	218270	2	AC079499 Mus muscu
20	28	0.9	223344	2	AC106193 Rattus no
21	26	0.8	39	6	AX684309 Sequence
22	25	0.8	37	6	AX684310 Sequence
23	25	0.8	37	6	AX684312 Sequence
24	24	0.8	38	6	AX684311 Sequence
25	24	0.8	110000	2	AC146908 Oryza sat
26	24	0.8	135406	8	AC136843 Oryza sat
27	24	0.8	144077	2	AC135398 Oryza sat
28	24	0.8	179591	2	AP004805 Oryza sat
29	24	0.8	252463	2	AC139169 Oryza sat
30	23	0.7	227555	10	AC108416 Mus muscu
31	23	0.7	231864	2	AC097246 Rattus no
32	23	0.7	240461	2	AC111669 Rattus no
33	23	0.7	249541	2	AC095672 Rattus no
34	23	0.7	249774	2	AC126196 Rattus no
35	22	0.7	114098	9	AL162594 Human DNA
36	22	0.7	156094	9	AC090198 Homo sapi
37	22	0.7	183778	2	AC011004 Homo sapi
38	22	0.7	184548	2	AC023551 Homo sapi
39	22	0.7	193882	9	AF313041 Homo sapi
40	22	0.7	199437	2	AC091575 Homo sapi
41	22	0.7	207069	9	AC090922 Homo sapi
42	22	0.7	207190	9	AF295017 Homo sapi
43	22	0.7	209296	5	AL954150 Zebrafish
44	22	0.7	211018	5	BX004964 Zebrafish
45	22	0.7	218270	2	AC079499 Mus muscu

ALIGNMENTS

RESULT 1						
AX417214						
LOCUS	AX417214	Sequence 23 from Patent WO0226780.	3108 bp	DNA	linear	PAT 14-JUN-2002
DEFINITION	AX417214					
ACCESSION	AX417214					
VERSION	AX417214.1	GI:21449784				
KEYWORDS	.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	Reed,J.C., Godzik,A., Chu,Z.L., Pawlowski,K., Fiorentino,L.,				
AUTHORS	Ariza,M.E. and Stehlik,C.					
TITLE	Paad domain-containing polypeptides, encoding nucleic acids, and					

JOURNAL	methods of use	
Patent: WO 0226780-A 23 04-APR-2002;	Patent: WO 0226780-A 23 04-APR-2002;	
The Burnham Institute (US)	The Burnham Institute (US)	
FEATURES	Location/Qualifiers	
source	1..3108	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	1..3108	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD35284.1"	
	/db_xref="GI:21449785"	
	/db_xref="REMBL:CAD35284"	
	/translation="MLRTAGDGLCRSLSTYIEELBAVELKKPKFLVLTATBELGKIP WSMEKAGLEMAQLLI THFPEEAWRIALSTFERINRKLWEROREDIYRDPQETV RUYVRKFLMEDRNARLGEVCNLSHRVTRLLLVKEHNSPMQVQQLLDTRGRHRTV GHQASPIKETLFEEDDEERPEPRVTRQAGAGIKGSMIAHKWMDWADGKLFQGRF YHPIINCRMNQSAETCSQMDLIFSCWPEPSAPLOELIRVERLLFIIDGDELKPSF ILGQGWKCEKPTLELLNSLRLLPELSLITRPTALEKHLRLEHPRHVE ILGFSAERKETFYFYFHNAEQGVFNVRDNEPLFTWCFVPLVWVCTCLOQOLE GGGLRKSRTTAVMYLLSLMQPKGAPRLQPPNORGLCSLAADGLWNOKLPE EODLRKHGLDGEDVSAFLNMNLFQKDINCERYVSIPIHLSFQEPFPAAMYIILDEGGA FQDQVTRLLTETYSERSFLATLSRFLFGLNESTRSHLEKSLCWSVPHIKMDLQ WFSQASQDSTLQOGLSEFFELYEIOEEEFIOALSHFQVI VVSNFASKMEHWS FCLKRSQAIVLHYGATSDGEDRARCAGHAHTLVQLRPERVTLIDAYSEHLAA LCTNPNLIELSVRNALSGRGVKKLCOGLRHENCKLQRLKRCRISISACEDLSAAL IANKNLRMDLSGNGVRPGMWMLCEGLRHPQCRLOMLRKCQLESACDEMA5VLG TNPHLVELDLTGNALDGLRLCOGLRHPVCRRLTWLKI CRLTAAACDELASTLSV NOSLRDLUSNELDGLVLLCEGLRHPTKLOTLRIGICRLGSAACBGLSVVLQAN HNLRLEULSFNDLGDWGMWLLAEGLQHPACRLQKWLSDCGLTAKACENLYFTLGINQ TLTDLVLTNNALGDTVRLCKRLSHPGCKLRVLFGMDLNKMTSHSLAALRVTKPY LDIGC"	
ORIGIN		
Query Match	100.0%; Score 3108; DB 6; Length 3108;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 3108; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGCTACGAACCGCAGGACGGCCCTCTGCGCCTGTCCACCTACTTGGAGAACTC 60
DB	1	ATGCTACGAACCGCAGGACGGCCCTCTGCGCCTGTCCACCTACTTGGAGAACTC 60
QY	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTATACCTGGGACCGGCACAGAGCTGGAGAA 120
DB	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTATACCTGGGACCGGCACAGAGCTGGAGAA 120
QY	121	GGCAAGATCCCTGGGGAAGCATGAGAAAGCGCGGTGCCCTGGAAATGGCCAGCTGCTC 180
DB	121	GGCAAGATCCCTGGGGAAGCATGAGAAAGCGCGGTGCCCTGGAAATGGCCAGCTGCTC 180
QY	181	ATCACCCACTTCGGGCCAGAGAGCGCTGGAGTTGGCTCTCAGCACCTTTGACGGGATA 240
DB	181	ATCACCCACTTCGGGCCAGAGAGCGCTGGAGTTGGCTCTCAGCACCTTTGACGGGATA 240
QY	241	AACAGGAAGGACTCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAGGGATCCCCAGGAA 300
DB	241	AACAGGAAGGACTCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAGGGATCCCCAGGAA 300
QY	301	ACCTACAGGACTATGTCCGAGGAAAATTCCGGCTCATGGAAGACCGCAATGCGCGCTA 360
DB	301	ACCTACAGGACTATGTCCGAGGAAAATTCCGGCTCATGGAAGACCGCAATGCGCGCTA 360
QY	361	GGGGAATGTCAACTCAGCACCGGTACACCGGGCTCCTGCTGGTGAAGGAGACTCA 420
DB	361	GGGGAATGTGTCAACTCAGCACCGGTACACCGGGCTCCTGCTGGTGAAGGAGACTCA 420
QY	421	AACCCCATGAGGCTCCAGCAGAGCTTCTGGACACAGCCGGGGACACGCGAGGACCGTG 480
DB	421	AACCCCATGAGGCTCCAGCAGAGCTTCTGGACACAGCCGGGGACACGCGAGGACCGTG 480
QY	481	GGACACAGGCTAGGCCCATCAAGATAGAGACCCCTCTTTGACCCACAGAGCGGCCCC 540
DB	481	GGACACAGGCTAGGCCCATCAAGATAGAGACCCCTCTTTGACCCACAGAGCGGCCCC 540

1621 AACGAGGAGACGAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTCTCTCGCGCACATC 1680
1621 AACGAGGAGACGAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTCTCTCGCGCACATC 1680
1681 AAGATGAGCCTGTTGAGTGGATCAAAAGCAAAAGCTCAGAGGAGCGGTCCACCTGCGAG 1740
1681 AAGATGAGCCTGTTGAGTGGATCAAAAGCAAAAGCTCAGAGGAGCGGTCCACCTGCGAG 1740
1741 CAGGCTCCTTGGAGTCTTCAAGCTGCTGTACGAGATCCAGAGAGGAGTTCATCCAG 1800
1741 CAGGCTCCTTGGAGTCTTCAAGCTGCTGTACGAGATCCAGAGAGGAGTTCATCCAG 1800
1801 CAGGCTCCTGAGCCTTCCAGCTGATCGTGGTCAGCAACATTCGCTCCAAAGATGAGGAC 1860
1801 CAGGCTCCTGAGCCTTCCAGCTGATCGTGGTCAGCAACATTCGCTCCAAAGATGAGGAC 1860
1861 ATGGTCTCCTGTTCTGTAAGCGCTGACAGAGCGCCAGGTCGTGCACTTGTATGGC 1920
1861 ATGGTCTCCTGTTCTGTAAGCGCTGACAGAGCGCCAGGTCGTGCACTTGTATGGC 1920
1921 GCCACTACAGCGGAGCGGGAAGACCGCGAGGTGCTCCGAGGAGCGCACAGCTG 1980
1921 GCCACTACAGCGGAGCGGGAAGACCGCGAGGTGCTCCGAGGAGCGCACAGCTG 1980
1981 TTGGTGCAGCTCAGACCCAGAGAGGACCGTCTCTGCTGGAGCGCTACAGTGAACATCTGGCA 2040
1981 TTGGTGCAGCTCAGACCCAGAGAGGACCGTCTCTGCTGGAGCGCTACAGTGAACATCTGGCA 2040
2041 GCGGCTCTGTGACCAATCCAACTGATAGAGTGTCTGTACCGAAATGCCCTGGGC 2100
2041 GCGGCTCTGTGACCAATCCAACTGATAGAGTGTCTGTACCGAAATGCCCTGGGC 2100
2101 AGCGGGGGGTGAAGTCTCTGCTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
2101 AGCGGGGGGTGAAGTCTCTGCTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
2161 CTGAGGCTGAAGAGGTGCGCGATCTCCAGCTCAGGAGTCTCTCTGAGCTCTC 2220
2161 CTGAGGCTGAAGAGGTGCGCGATCTCCAGCTCAGGAGTCTCTCTGAGCTCTC 2220
2221 ATAGCCAATAAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGTGGATTCACAGC 2280
2221 ATAGCCAATAAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGTGGATTCACAGC 2280
2281 ATGATGCTGCTTTGAGGAGGCTGCGGATCCCACTGAGGAGTCTGAGTGTGAGT 2340
2281 ATGATGCTGCTTTGAGGAGGCTGCGGATCCCACTGAGGAGTCTGAGTGTGAGT 2340
2341 AGGAAAGTGTACGCTGAGTCCGGGCTTGTCTGAGGAGTGGCTTCTGCTCGGCAACCAAC 2400
2341 AGGAAAGTGTACGCTGAGTCCGGGCTTGTCTGAGGAGTGGCTTCTGCTCGGCAACCAAC 2400
2401 CCACATCTGGTTGAGTGGACCTGACAGAAATGCACTGAGGAGTGGCTTCTGCTCGGCAACCAAC 2460
2401 CCACATCTGGTTGAGTGGACCTGACAGAAATGCACTGAGGAGTGGCTTCTGCTCGGCAACCAAC 2460
2461 CTATGCCAGGAGTGGAGGACCCAGTCTGACACTACGAGTCTGCTGAGTGTGAGTCTG 2520
2461 CTATGCCAGGAGTGGAGGACCCAGTCTGACACTACGAGTCTGCTGAGTGTGAGTCTG 2520
2521 CGCCTCACTGCTGCTGTGACAGGCTGGCTCAACTCTCAGTGTGTGAACCCAGAGCCTG 2580
2521 CGCCTCACTGCTGCTGTGACAGGCTGGCTCAACTCTCAGTGTGTGAACCCAGAGCCTG 2580
2581 AGAGAGTGGAACTGAGCTGAGCTGAGGAGCCTCGGGTGTGCTGTGTGAGT 2640
2581 AGAGAGTGGAACTGAGCTGAGCTGAGGAGCCTCGGGTGTGCTGTGTGAGT 2640
2641 GGCCTCAGGCATCCACGCTGAGCTCCAGACCTGCGGTTGGGATCTGCGGCTGGGC 2700
2641 GGCCTCAGGCATCCACGCTGAGCTCCAGACCTGCGGTTGGGATCTGCGGCTGGGC 2700

2701 TCTGCCGCTGTGAGGGTCTTCTGCTGCTGCCAGGCAACCAACCTCCGGGAGCTG 2760
2701 TCTGCCGCTGTGAGGGTCTTCTGCTGCTGCCAGGCAACCAACCTCCGGGAGCTG 2760
2761 GACTTGAAGTTTCAACGACTGGGAGCTGGGGCTGTGGTTGCTGGCTGAGGGCTGCAA 2820
2761 GACTTGAAGTTTCAACGACTGGGAGCTGGGGCTGTGGTTGCTGGCTGAGGGCTGCAA 2820
2821 CATCCCGCTGCAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCTCACAGCCAGGCT 2880
2821 CATCCCGCTGCAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCTCACAGCCAGGCT 2880
2881 TGTGAGAATCTTTACTTTCACCTGGGATCAACAGAGCTTTCACCGACCTTTACCTGACC 2940
2881 TGTGAGAATCTTTACTTTCACCTGGGATCAACAGAGCTTTCACCGACCTTTACCTGACC 2940
2941 AACACGCGCTAGGGGACACAGGTGTCCGACTGTCTTTCGAAAGCGGTGAGCCATCTGCGC 3000
2941 AACACGCGCTAGGGGACACAGGTGTCCGACTGTCTTTCGAAAGCGGTGAGCCATCTGCGC 3000
3001 TGCAAACTCCGAGTCTCTGCTTATTTGGGATGGACCTGAATAAATGACCCACAGTAGG 3060
3001 TGCAAACTCCGAGTCTCTGCTTATTTGGGATGGACCTGAATAAATGACCCACAGTAGG 3060
3061 TTGGCAGCGCTTCGAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108
3061 TTGGCAGCGCTTCGAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108

RESULT 2
AX459869
LOCUS AX459869 3300 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 12 from Patent WO0240668.
ACCESSION AX459869
VERSION AX459869.1 GI:21725643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Tschoep, J. and Martinon, F.
TITLE Proteins and dna sequences underlying these proteins used for treating inflammations
JOURNAL Patent: WO 0240668-A 12 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source Location/Qualifiers
1..3300
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP3/PY5.dna"

ORIGIN
Query Match 98.4%; Score 3057; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
Db 1 ATGTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
QY 61 GAGGCTGTGGAACTCAAGAAGTTCAAGTTATACCTGGGAGCCGACACAGAGCTGGAGAA 120
Db 61 GAGGCTGTGGAACTCAAGAAGTTCAAGTTATACCTGGGAGCCGACACAGAGCTGGAGAA 120
QY 121 GCGAAGATCCCTCGGGGAAGCATGAGAAAGCGCGCTGCCCTGGAATGSCCAGCTGTCT 180
Db 121 GCGAAGATCCCTCGGGGAAGCATGAGAAAGCGCGCTGCCCTGGAATGSCCAGCTGTCT 180
QY 181 ATCACCCACTTGGGCCACAGAGAGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
Db 181 ATCACCCACTTGGGCCACAGAGAGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240

QY 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGGTGAAGGATCCCCAGGAA 300
DB 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGGTGAAGGATCCCCAGGAA 300
QY 301 ACCTACAGGAGCTATGTCCGAGAGAAATTCGGGTCTATGGAAGACCGCAATGCGGCTTA 360
DB 301 ACCTACAGGAGCTATGTCCGAGGAAATTCGGGTCTATGGAAGACCGCAATGCGGCTTA 360
QY 361 GGGGAATGTGTCAACCTCAGCACCGGTACACCGGCTCCGTGGTGAAGGACACTCA 420
DB 361 GGGGAATGTGTCAACCTCAGCACCGGTACACCGGCTCCGTGGTGAAGGACACTCA 420
QY 421 AACCCCATGACAGGTCCAGCAGCAGGTTCTTGACACAGCCCGGGACACGCGAGGACCGTG 480
DB 421 AACCCCATGACAGGTCCAGCAGCAGGTTCTTGACACAGCCCGGGACACGCGAGGACCGTG 480
QY 481 GGACACAGAGGTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAGAGGCGGCC 540
DB 481 GGACACAGAGGTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAGAGGCGGCC 540
QY 541 GAGCACCGCACCGGTGTATGCAAGCGCGGAGGATAGGCAAGTCCATGCTGCA 600
DB 541 GAGCACCGCACCGGTGTATGCAAGCGCGGAGGATAGGCAAGTCCATGCTGCA 600
QY 601 CACAAGGTGTATGCTGAGCTGGGCGGACGGGAAGCTCTTCCAAGGCAGATTTGATTCTC 660
DB 601 CACAAGGTGTATGCTGAGCTGGGCGGACGGGAAGCTCTTCCAAGGCAGATTTGATTCTC 660
QY 661 TTCTACATCACTGAGGAGATGAACAGAGTGGCAGCGAATGAGGATGCAAGCTC 720
DB 661 TTCTACATCACTGAGGAGATGAACAGAGTGGCAGCGAATGAGGATGCAAGCTC 720
QY 721 ATCTTACGCTCTGCGCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
DB 721 ATCTTACGCTCTGCGCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
QY 781 CGCTCTCTTTTCATCAGCGCTTCGATGAGCTCAAGCCTCTTTCCAGATCCTCAG 840
DB 781 CGCTCTCTTTTCATCAGCGCTTCGATGAGCTCAAGCCTCTTTCCAGATCCTCAG 840
QY 841 GGACCTGTGCTCTGCTGGAGAGAAACGGCCGACGAGCTGCTTTAAAGCTTA 900
DB 841 GGACCTGTGCTCTGCTGGAGAGAAACGGCCGACGAGCTGCTTTAAAGCTTA 900
QY 901 ATTCGGAAGAGCTGCTCCCTGAGCTATCTTGTCTCATCACACAGCGCCACGCTTG 960
DB 901 ATTCGGAAGAGCTGCTCCCTGAGCTATCTTGTCTCATCACACAGCGCCACGCTTG 960
QY 961 GAGAAGCTCCACCTGCTGAGCACCCAGGATGTGGAGATCCTGGGCTTCTCTGAG 1020
DB 961 GAGAAGCTCCACCTGCTGAGCACCCAGGATGTGGAGATCCTGGGCTTCTCTGAG 1020
QY 1021 GCAGAAAGGAGTAATCTTCAAGTATTTCAACATGACAGAGCGGCGCAAGTC 1080
DB 1021 GCAGAAAGGAGTAATCTTCAAGTATTTCAACATGACAGAGCGGCGCAAGTC 1080
QY 1081 TTCAATTAAGTGAGGACCAACGAGCTCTTTCAACATGCTCTGCTCCCTGCTGTC 1140
DB 1081 TTCAATTAAGTGAGGACCAACGAGCTCTTTCAACATGCTCTGCTCCCTGCTGTC 1140
QY 1141 TGGTGTGTGTACTGCTCCAGCAGCTGAGGAGTGGGGGCTTGTGAGACAGAC 1200
DB 1141 TGGTGTGTGTACTGCTCCAGCAGCTGAGGAGTGGGGGCTTGTGAGACAGAC 1200
QY 1201 TCAGAGCCACCTGAGTGTATGCTCTACCTGCTGAGTCTGATGCAACCAAGCC 1260
DB 1201 TCAGAGCCACCTGAGTGTATGCTCTACCTGCTGAGTCTGATGCAACCAAGCC 1260
QY 1261 GGGGCCCCGCGCTCCAGCCCCCAGAGAGGTTGTGCTCTTGGCGGAGAT 1320
DB 1261 GGGGCCCCGCGCTCCAGCCCCCAGAGAGGTTGTGCTCTTGGCGGAGAT 1320

QY 1321 GGGCTCTGGAATCAGAAAATCTATTTAGGAGCAGGACCTCCGGAAGCAGGCTAGAC 1380
DB 1321 GGGCTCTGGAATCAGAAAATCTATTTAGGAGCAGGACCTCCGGAAGCAGGCTAGAC 1380
QY 1381 GGGGAAGACGTCTCTGCTTCTCAACATGAACATCTTCCAGAAAGGACATCAACTGTGAG 1440
DB 1381 GGGGAAGACGTCTCTGCTTCTCAACATGAACATCTTCCAGAAAGGACATCAACTGTGAG 1440
QY 1441 AGGTACTACAGCTTCACTCCTTCTGAGTTTCCAGAAATCTTGTGAGCTATGTACTATATC 1500
DB 1441 AGGTACTACAGCTTCACTCCTTCTGAGTTTCCAGAAATCTTGTGAGCTATGTACTATATC 1500
QY 1501 CTGACAGAGGGGAGGCGGGCCAGCCACACAGCAGGACGTGACAGGCTGTGACCGAG 1560
DB 1501 CTGACAGAGGGGAGGCGGGCCAGCCACACAGCAGGACGTGACAGGCTGTGACCGAG 1560
QY 1561 TACCGCTTTTCTGAAAGAGCTTCTGGCACTCACCAGCCGCTTCTGTTTGGACTCTG 1620
DB 1561 TACCGCTTTTCTGAAAGAGCTTCTGGCACTCACCAGCCGCTTCTGTTTGGACTCTG 1620
QY 1621 AACGAGGAGACGAGGACCACTCGGAGAGAGTCTCTGCTGGAAGGTCTCGCCGACATC 1680
DB 1621 AACGAGGAGACGAGGACCACTCGGAGAGAGTCTCTGCTGGAAGGTCTCGCCGACATC 1680
QY 1681 AAGATGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGCGACGGCTCCACCTGCGAG 1740
DB 1681 AAGATGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGCGACGGCTCCACCTGCGAG 1740
QY 1741 CAGGCTCTTGGAGTTTCTTCCAGCTGTTGACAGATCCAGGAGGAGGATTTATCCAG 1800
DB 1741 CAGGCTCTTGGAGTTTCTTCCAGCTGTTGACAGATCCAGGAGGAGGATTTATCCAG 1800
QY 1801 CAGGCTCTGAGCCACTTCCAGGTGATCGTGTGAGCAACTTGTCTTCAAGATGGAGCAC 1860
DB 1801 CAGGCTCTGAGCCACTTCCAGGTGATCGTGTGAGCAACTTGTCTTCAAGATGGAGCAC 1860
QY 1861 ATGTCTCTGCTGTTCTGAAAGCGTGCAGGAGCGCCAGGTGCTGCACTTGTATGCG 1920
DB 1861 ATGTCTCTGCTGTTCTGAAAGCGTGCAGGAGCGCCAGGTGCTGCACTTGTATGCG 1920
QY 1921 GCCACCTACAGCGGAGCGGAGAGACCGCGAGGTGCTCCGAGGAGCGCACACGCTG 1980
DB 1921 GCCACCTACAGCGGAGCGGAGAGACCGCGAGGTGCTCCGAGGAGCGCACACGCTG 1980
QY 1981 TTGTGTGAGCTCAGACACGAGGACCGTCTGCTGAGAGCGCTACAGTGAACATCTGGCA 2040
DB 1981 TTGTGTGAGCTCAGACACGAGGACCGTCTGCTGAGAGCGCTACAGTGAACATCTGGCA 2040
QY 2041 GCGGCTCTGTGCAACAAATCCAAACCTGATAGAGTGTCTGTATCCGAAATGCCCCGGGC 2100
DB 2041 GCGGCTCTGTGCAACAAATCCAAACCTGATAGAGTGTCTGTATCCGAAATGCCCCGGGC 2100
QY 2101 AGCCGGGGGTGAGCTGCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
DB 2101 AGCCGGGGGTGAGCTGCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
QY 2161 CTGAGGCTGAAGAGTGCCTCATCTCCAGCTCAGCTGCGAGGACCTCTCTGCACTCTC 2220
DB 2161 CTGAGGCTGAAGAGTGCCTCATCTCCAGCTCAGCTGCGAGGACCTCTCTGCACTCTC 2220
QY 2221 ATAGCCAAATGAATTTGCAAGGATGATCTCAGTGGCAACCGGCTTGGATTTCCAGGC 2280
DB 2221 ATAGCCAAATGAATTTGCAAGGATGATCTCAGTGGCAACCGGCTTGGATTTCCAGGC 2280
QY 2281 ATGATGCTGCTTTCGAGGGGCTCGGCATCCCCAGTGCAGCTGAGATGATTCAAGTTG 2340
DB 2281 ATGATGCTGCTTTCGAGGGGCTCGGCATCCCCAGTGCAGCTGAGATGATTCAAGTTG 2340
QY 2341 AGGAAGTGTACGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTCTCGGCAACCAAC 2400
DB 2341 AGGAAGTGTACGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTCTCGGCAACCAAC 2400
QY 2401 CCACATCTGTTGAGTTGGACCTCAGAGGAATGCACTGGAGGATTTGGGCTCGAGTTA 2460

Db	2401	CCACATCTGGTTGAGTTCGAGCTGACAGAAATGCACTGGAGGATTTGGGCTGAGGTTA	2460
Qy	2461	CTATGCGAGGAGCTGAGSACCCAGTCTGCGAGACTACGAGCTTTGTGCTGAAGATCTGC	2520
Db	2461	CTATGCGAGGAGCTGAGSACCCAGTCTGCGAGACTACGAGCTTTGTGCTGAAGATCTGC	2520
Qy	2521	CGCCTCACTGCTGCTGCTGCTGACAGCTGGGCTCAA	2580
Db	2521	CGCCTCACTGCTGCTGCTGACAGCTGGGCTCAA	2580
Qy	2581	AGAGAGCTGGACCTGAGGCTGAATGAGCTGGGGGACCTCGGGGTCTGCTGCTGTGAG	2640
Db	2581	AGAGAGCTGGACCTGAGGCTGAATGAGCTGGGGGACCTCGGGGTCTGCTGCTGTGAG	2640
Qy	2641	GGCCTCAGGCATCCACGCTGACAGCTCCAGACCTCGGTTGGGCTGCGGGTGGGC	2700
Db	2641	GGCCTCAGGCATCCACGCTGACAGCTCCAGACCTCGGTTGGGCTGCGGGTGGGC	2700
Qy	2701	TCTGCGGCTGTGAGGGTCTTCTGCTGGTCTCCAGGGCAACCA	2760
Db	2701	TCTGCGGCTGTGAGGGTCTTCTGCTGGTCTCCAGGGCAACCA	2760
Qy	2761	GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTGGTCTGAGGGGCTGCAA	2820
Db	2761	GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTGGTCTGAGGGGCTGCAA	2820
Qy	2821	CATCCCGCTGCGAGACTCAGAACTGGCTGGATAGTGGCTGACAGCAAGGCT	2880
Db	2821	CATCCCGCTGCGAGACTCAGAACTGGCTGGATAGTGGCTGACAGCAAGGCT	2880
Qy	2881	TGTGAGAACTTTTACTTCAACCTGGGGATCAACAGACTTCAACGACCTTTACCTGACC	2940
Db	2881	TGTGAGAACTTTTACTTCAACCTGGGGATCAACAGACTTCAACGACCTTTACCTGACC	2940
Qy	2941	AACACGCGCTAGGGGACACAGGTGTCGAGCTTTTCAAGCGGCTGAGCCATCTGGC	3000
Db	2941	AACACGCGCTAGGGGACACAGGTGTCGAGCTTTTCAAGCGGCTGAGCCATCTGGC	3000
Qy	3001	TGCACACTCCGAGTCTCTGTTATTTGGGATGGACCTGAATAAATGACCCACAGT	3057
Db	3001	TGCACACTCCGAGTCTCTGTTATTTGGGATGGACCTGAATAAATGACCCACAGT	3057
RESULT 3			
AY116204			
LOCUS	AY116204	3731 bp mRNA linear PRI 21-MAY-2003	
DEFINITION	Homo sapiens monarch-1 mRNA, complete cds; alternatively spliced.		
ACCESSION	AY116204		
VERSION	AY116204.1	GI:21711820	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 3731)		
AUTHORS	Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.		
TITLE	Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes		
JOURNAL	J. Immunol. 170 (11), 5354-5358 (2003)		
PUBMED	12759408		
REFERENCE	2 (bases 1 to 3731)		
AUTHORS	Williams, K.L., Linhoff, M.W., Harton, J.A. and Ting, J.P.Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm		
Location/Qualifiers	Chapel Hill, NC 27599, USA		
FEATURES			
source	1..3731		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19q13.4"		

		/cell_line="U937"	
		221.._3409	
		/note="alternatively spliced"	
		/codon_start=1	
		/product="monarch-1"	
		/protein_id="AAM75142.1"	
		/db_xref="GI:21711821"	
		/translation="MLRTAGDGLCRSLTYLEAEVKKFKLVLTATGELGEGKIP	
		WGSKEKAGPLEMAOLLTHFGPEARLALSTFERINRKLWEGORELDVDTDPGG	
		PSSIGNQSTCLLESLVTPRXDPOETVRRKFLMEDNRARLCEGVNLSHRVTRL	
		LLVKEHNSPMQOOLLDTGRHARTVGHQASPIKIEFLFPDEPPEPPTVVMQGA	
		AGIQKSLAHKMLMDWADKLFQGRFDYLFYINCRMNQSAETECMSQDLIFSCWEP	
		APLDELIRVPERLLFIIDGDELKPSFDPQPCWELKREPTLELLINSLIKKLLP	
		ELSLITTRPTALEKHLRHLPRHVEILGFSEARKEKYFYKFINAAGQGVNVR	
		DNEPLTFMCFVPLNCVVCTCLOQLEGGLRQTSRTTAYVMLYLLSLMQPKGAP	
		RLQPPNORGLCSLAADGLWNQKILFERQDLRKHGLDGEDYSAFNNMIFQKINCER	
		YYSFHSUFQFFAAMYILDEGEGAGDQDVTKLLTEYAFSEFSFALTSRFLFGL	
		LINESTRHLEKSLCKWPSHIKMDLLOWISKAQSDSTLQOQSLFFSYLIEBEE	
		FIQALSHFOVIVVSNIAKSMHVSFCLKRCRSAQVHLHYGATYSADGEDRACSA	
		GAHTLLVOLRPERTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRVKLLCQGLRH	
		PNCQLNLRKRCRISSACEDLSAALANKNLTRMDLSGNGVGPFGMWMLCEGRHP	
		QCRLQMIOLRKQLESAGCEMASVLGTNPHVELDLTGNALEDGLGLLCCQGLRHP	
		CLRLTMUKICRLTAACDELASLTVNSQSLRELDLSLNEGLDGLGVLLCEGLRPTC	
		KLQTLRLGICRLGSAACGLSVLQAHNRLDELDFNDLGDWGLWLLAEGLQHAPCR	
		LQKLWDSGLTKAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPGCKL	
		RVLWLFQMDLKNMTHSLAALRVTKPYLDIGC"	
		869.._892	
		/note="Region: P-loop"	
		2069.._3406	
		/note="Region: leucine rich repeat"	
		221.._289	
		/note="Region: pyrin domain"	
		misc_feature	
		misc_feature	
		misc_feature	
		ORIGIN	
		Query Match 90.7%; Score 2820; DB 9; Length 3731;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2820; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	289	GATCCCCAGGAAACCTACAGGACTATGTCCGACAGAAATTCGGCTCATGAAGACCCG	348
Db	590	GATCCCCAGGAAACCTACAGGACTATGTCCGACAGAAATTCGGCTCATGAAGACCCG	649
Qy	349	AATGCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCTCTGCTGGTG	408
Db	650	AATGCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCTCTGCTGGTG	709
Qy	409	AAGGAGACTCAAAACCCATGAGTCCAGAGAGCTTCTGGACACAGGCGGGGACAC	468
Db	710	AAGGAGACTCAAAACCCATGAGTCCAGAGAGCTTCTGGACACAGGCGGGGACAC	769
Qy	469	CGGAGGCGCTGGGACACAGGCTAGCCCATCAAGATAGACACCTCTTTGAGCCAGAC	528
Db	770	CGGAGGCGCTGGGACACAGGCTAGCCCATCAAGATAGACACCTCTTTGAGCCAGAC	829
Qy	529	GAGGAGCGCCCGGAGCCACCGCGCACCGTGGTTCATGAAGCGCGGAGGATAGGCAAG	588
Db	830	GAGGAGCGCCCGGAGCCACCGCGCACCGTGGTTCATGAAGCGCGGAGGATAGGCAAG	889
Qy	589	TCCATGCTGGACACAAAGGTGATGCTGGGCGGACGGGAAGCTCTTCCAAAGGACA	648
Db	890	TCCATGCTGGACACAAAGGTGATGCTGGGCGGACGGGAAGCTCTTCCAAAGGACA	949
Qy	649	TTTGATTTATCTTCTTACATCAATGAGGAGATGAACAGAGTGCACGGAATGCAGC	708
Db	950	TTTGATTTATCTTCTTACATCAATGAGGAGATGAACAGAGTGCACGGAATGCAGC	1009
Qy	709	ATGCAAGACCTCATCTTCAGCTGCTGGCTGAGCCCGACCGCCCTCTCCAGAGCTCATC	768
Db	1010	ATGCAAGACCTCATCTTCAGCTGCTGGCTGAGCCCGACCGCCCTCTCCAGAGCTCATC	1069
Qy	769	CGAGTTCCGAGCGCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCTTTC	828
Db	1070	CGAGTTCCGAGCGCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCTTTC	1129

QY 829 CACGATCTCAGGACCCCTGTGTGCTCTGCTGGAGAGAGAAACGGCCACCGAGCTGCTT 888
DB CACGATCTCAGGACCCCTGTGTGCTCTGCTGGAGAGAGAAACGGCCACCGAGCTGCTT 1189
QY 889 CTTAACAGCTTAATTCGGAGAGAGCTGCTCCTGAGCTATCTTTGCTCATCACACAGG 948
DB CTTAACAGCTTAATTCGGAGAGAGCTGCTCCTGAGCTATCTTTGCTCATCACACAGG 1249
QY 949 CCCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCCGAGCATGTGGAGATCTG 1008
DB CCCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCCGAGCATGTGGAGATCTG 1309
QY 1009 GGCTTCTCTGAGGAGAGAGAGAAATTAATTTCAAGTATTTTCAATATTCAGAGAGAG 1068
DB GGCTTCTCTGAGGAGAGAGAGAAATTAATTTCAAGTATTTTCAATATTCAGAGAGAG 1369
QY 1069 GCGGGCCAAAGTCTTCAATTAAGTGGAGACCAAGAGCTCTCTTCAACATGTCTTGGTC 1128
DB GCGGGCCAAAGTCTTCAATTAAGTGGAGACCAAGAGCTCTCTTCAACATGTCTTGGTC 1429
QY 1129 CCCCTGTGTGCTGGGTGGTGTGTACCTGCTCCAGCAGAGAGCTGGAGGGTGGGGGCTG 1188
DB CCCCTGTGTGCTGGGTGGTGTGTACCTGCTCCAGCAGAGAGCTGGAGGGTGGGGGCTG 1489
QY 1189 TTGAGACAGAGCTCCAGAGACCACTGCGAGTGTACATGCTCTACCTGCTGAGTCTGATG 1248
DB TTGAGACAGAGCTCCAGAGACCACTGCGAGTGTACATGCTCTACCTGCTGAGTCTGATG 1549
QY 1249 CAACCCAAAGCGGGGGCCCGGCTCCAGCCCGCCACCCACAGAGAGGGTGTGCTCC 1308
DB CAACCCAAAGCGGGGGCCCGGCTCCAGCCCGCCACCCACAGAGAGGGTGTGCTCC 1609
QY 1309 TTGCGCGCAGATGGGCTCTGGAATCAGAAATCTATTTTGAAGGAGAGAGCTCCGGAAG 1368
DB TTGCGCGCAGATGGGCTCTGGAATCAGAAATCTATTTTGAAGGAGAGAGCTCCGGAAG 1669
QY 1369 CAGGGCTTACGCGGGAGAGCTCTGCTTCTCTCAACATGAACATCTTCCAGAGAGAC 1428
DB CAGGGCTTACGCGGGAGAGCTCTGCTTCTCTCAACATGAACATCTTCCAGAGAGAC 1729
QY 1429 ATCAACTGTGAGAGTACTACAGTCTCATCCACTTGTAGTTTCCAGGAATCTTTTGCAGCT 1488
DB ATCAACTGTGAGAGTACTACAGTCTCATCCACTTGTAGTTTCCAGGAATCTTTTGCAGCT 1789
QY 1489 ATGTACTATATCTTGAAGGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGGG 1548
DB ATGTACTATATCTTGAAGGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGGG 1849
QY 1549 CTGTTGACCGAGTACGGTCTTCTGAAGAGCTTCTGGACCTTCCAGAGAGCTTCCGCTG 1608
DB CTGTTGACCGAGTACGGTCTTCTGAAGAGCTTCTGGACCTTCCAGAGAGCTTCCGCTG 1909
QY 1609 TTTGGACTCTGTAACAGAGAGAGACAGGAGCCACTGAGAGAGAGTCTCTGCTGGAAGTCT 1668
DB TTTGGACTCTGTAACAGAGAGAGACAGGAGCCACTGAGAGAGAGTCTCTGCTGGAAGTCT 1969
QY 1669 TCGCCGCACATCAAGATGAGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGAGAGAG 1728
DB TCGCCGCACATCAAGATGAGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGAGAGAG 2029
QY 1729 TCCACCTGACAGAGGCTCTGAGGATCTTCAAGTGTCTTCAAGTGTCTTCAAGTGTCTTCA 1788
DB TCCACCTGACAGAGGCTCTGAGGATCTTCAAGTGTCTTCAAGTGTCTTCAAGTGTCTTCA 2089
QY 1789 GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGGTGTCTGAGTGTCTGAGTGTCTGAGTGT 1848
DB GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGGTGTCTGAGTGTCTGAGTGTCTGAGTGT 2149
QY 1849 AAGATGAGACATGTGCTCTGCTTCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1908
DB AAGATGAGACATGTGCTCTGCTTCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 2209

QY 1909 CACTTGTATGCGCCACCTACAGCGCGGACCGGGAAGACCGCGAGGTGCTCCGAGGA 1968
DB CACTTGTATGCGCCACCTACAGCGCGGACCGGGAAGACCGCGAGGTGCTCCGAGGA 2269
QY 1969 GCGCACACGCTGTTGGTGCAGCTCAGACAGAGAGGACCGTTCGTGCGAGCGCTACAGT 2028
DB GCGCACACGCTGTTGGTGCAGCTCAGACAGAGAGGACCGTTCGTGCGAGCGCTACAGT 2329
QY 2029 GAACATCTGGCAGCGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTCTGTAACCGA 2088
DB GAACATCTGGCAGCGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTCTGTAACCGA 2389
QY 2089 AATCCCTGGGACCGCGGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCAACTGC 2148
DB AATCCCTGGGACCGCGGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCAACTGC 2449
QY 2149 AACTCTCAGAACCTGAGGCTGAGAGGTGCGGATCTCCAGCTCAGCTCAGCTGCGAGGACCTC 2208
DB AACTCTCAGAACCTGAGGCTGAGAGGTGCGGATCTCCAGCTCAGCTCAGCTGCGAGGACCTC 2509
QY 2209 TCTCAGCTCTCATAGCCAATTAAGAAATTTGACAAAGGATGATCTCAGTGGCAACGGCGTT 2268
DB TCTCAGCTCTCATAGCCAATTAAGAAATTTGACAAAGGATGATCTCAGTGGCAACGGCGTT 2569
QY 2269 GATTTCCGAGCATGTGCTGTTTGGAGGGCTGCGGATCCCGATCCCGAGTGCAGCTGCAG 2328
DB GATTTCCGAGCATGTGCTGTTTGGAGGGCTGCGGATCCCGATCCCGAGTGCAGCTGCAG 2629
QY 2329 ATGATTCAGTTGAGGAGTGTGCTGAGTGTGAGTGTGAGCTGACGAGAAATGCACTGGAGATTG 2388
DB ATGATTCAGTTGAGGAGTGTGCTGAGTGTGAGTGTGAGCTGACGAGAAATGCACTGGAGATTG 2689
QY 2389 CTGCGGCAACCAACCATCTGTTGAGTGTGAGCTGACGAGAAATGCACTGGAGATTG 2448
DB CTGCGGCAACCAACCATCTGTTGAGTGTGAGCTGACGAGAAATGCACTGGAGATTG 2749
QY 2449 GGCCTGAGGTTATATGCGCAGGAGCTGAGGACCCAGTCTGACAGACTACGAGCTTTGFG 2508
DB GGCCTGAGGTTATATGCGCAGGAGCTGAGGACCCAGTCTGACAGACTACGAGCTTTGFG 2809
QY 2509 CTGAAGATCTGCGGCTCAGTGTGCTGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2568
DB CTGAAGATCTGCGGCTCAGTGTGCTGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2869
QY 2569 AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2628
DB AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2929
QY 2629 CTGCTGTGAGGCGCTCAGGACCTCCAGCTGCAAGCTCCAGACCTTGGGCTGCGGATC 2688
DB CTGCTGTGAGGCGCTCAGGACCTCCAGCTGCAAGCTCCAGACCTTGGGCTGCGGATC 2989
QY 2689 TGCGGCTGCGCTGCGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 2748
DB TGCGGCTGCGCTGCGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 3049
QY 2749 CTGCGGAGCTGAGCTTGAATTTCAACGAGCTGAGGAGCTGAGGCTGAGGCTGAGGCT 2808
DB CTGCGGAGCTGAGCTTGAATTTCAACGAGCTGAGGAGCTGAGGCTGAGGCTGAGGCT 3109
QY 2809 GAGGGCTGCAACATCCCGCTGAGACTCAGAACTGAGGCTGAGGCTGAGGCTGAGGCT 2868
DB GAGGGCTGCAACATCCCGCTGAGACTCAGAACTGAGGCTGAGGCTGAGGCTGAGGCT 3169
QY 2869 ACAGCCAGGCTTGTGAGATCTTTACTTCCACCTGGGATCAACAGACTTGAACGAC 2928
DB ACAGCCAGGCTTGTGAGATCTTTACTTCCACCTGGGATCAACAGACTTGAACGAC 3229
QY 2929 CTTTACCTGACCAACACCGCTTAGGGGACAGAGGTGTCGAGTGTTCGAGCGGCTG 2988
DB CTTTACCTGACCAACACCGCTTAGGGGACAGAGGTGTCGAGTGTTCGAGCGGCTG 3289
QY 2989 AGCCATCTGGCTGCAACCTCCGAGTCTCTGTTATTTGGGATGGAGCTGAATAATG 3048

Db 3290 AGCATCTGGCTGCAAACTCCGAGTCTCTGGTTATTGGATGGACCTGAATAAATG 3349
 QY 3049 ACCACAGTAGTGGGAGCGCTTCGAGTAAACAAACCTTATTGGACATGGCTGCTGA 3108
 Db 3350 ACCACAGTAGTGGGAGCGCTTCGAGTAAACAAACCTTATTGGACATGGCTGCTGA 3409

RESULT 4
 BC028069
 LOCUS
 DEFINITION Homo sapiens NACHT, leucine rich repeat and PYD containing 12, mRNA PRI 04-NOV-2003
 (cDNA clone MGC:40117 IMAGE:5212737), complete cds.
 ACCESSION BC028069
 VERSION
 KEYWORDS
 SOURCE MGC.
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 3507)
 Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hukley, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smalhus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 PUBMED
 22388257
 12477932

REFERENCE
 AUTHORS
 2 (bases 1 to 3507)
 Strausberg, R.
 Direct Submission
 Submitted (08-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgi.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@hgrl.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 62 Row: 9 Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21955153.

FEATURES
 source
 Location/Qualifiers
 1..3507
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:40117 IMAGE:5212737"
 /tissue_type="Blood, adult leukocytes"
 /clone_lib="NIH_MGC_118"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..3507
 /gene="NALP12"
 /note="synonyms: RNO2, PYPAF7"
 /db_xref="LocusID:91662"
 120..3305
 /codon_start=1
 /product="PYRIN-containing APAF1-like protein 7"
 /protein_id="AAH28069.1"
 /db_xref="GI:20380400"
 /db_xref="LocusID:91662"
 /translation="MLRTAGDGLCRSLTYLEELEAVELKFKYVLGTATLGEKIP
 WSGMERAGPLEMAQLLTHFGPEAWRLALSTFERINRKLWQRQEDLVDTDPGG
 PSSLGNGSTCLLESLVTPRKDPQETVYRRFRLMEDNARLGEVNLSHRYTRL
 LLVKGNSPMQVOQLDITGRGHARTVGHQASPIKLTLPEDPERPPPTVVMQGA
 AGICKMLAHKVMLDWADGKLFQGRFDYLFVINCENMQSATECMQDLIFCWPEPS
 AQLGLRVRPRLIFIDGPELAPSHDPQWPCWLCEWRKPELLELNSLRKLLP
 ELSLITTRTALKEKRLHRLHPRHEILFHPDSEARKEIFYKFHNASQAGQVFNVR
 RLEPFTMGCLVAVCTCLOQLGEGGLRLQRTSTRTAYMLYLUSLMQPKGAP
 RLPQPNRGICSLAADGKILFEEQDLRKHGLEDGEVSAFLNMNIQDKINCR
 YSFTHLSFQFAAMYILDEGGAGPDQDVTLLTEYAFSPRSFLATSRFLFGL
 LNESTHLEKSLCKWSPHMKDLOWIQSKAOSDSTLOQSLERFSCYLEIQEE
 FIQALSHPQVIVSNVSNASKHMSVSCFKRCSAQLVHLIYATYADGEDRCSA
 GAHTLVQLPRTVLDAISHLAALCTNKLRLDMSGNGVGFMMLLCGLRHP
 NCKLQNLRLKRISSEADLSAALIANKLRLDMSGNGVGFMMLLCGLRHP
 CLRLQILRLKQESACDEMASVLGNPHLVELDTGNALDELGLLCCGLRHPVC
 RLQTLWLKICRLTAACDELASTLSVNOSLRELDLSNDLGDMLWLLAEGLOHPCTCK
 LQTLRLGICRLGSAACGLSVLQANHLRELDLSNDLGDMLWLLAEGLOHPCTCK
 QKWLDDSGCLTKACACENDIFLTGINQTLIDYLLINNALGDTGVRLCKRLSHPGCKLR
 VLWLFMDLNKMTWTHSLAALRVTKPYLDIGC"
 153..395
 /note="PAAD DAPIN; Region: PAAD/DAPIN/Pyrin domain. This
 domain is predicted to contain 6 alpha helices and to have
 the same fold as the pfam00531 domain. This similarity may
 mean that this is a protein-protein interaction domain"
 /db_xref="CDD:pfam02758"
 753..1262
 /note="NACHT; Region: NACHT domain. This NTPase domain is
 found in apoptosis proteins as well as those involved in
 MHC transcription activation. This family is closely
 related to pfam00931"
 /db_xref="CDD:pfam05729"
 2268..3224
 /note="LRR RI; Region: Leucine-rich repeats (LRRs),
 ribonuclease inhibitor (RI)-like subfamily. LRRs are 20-29
 residue sequence motifs present in many proteins that
 participate in protein-protein interactions and have
 different functions and cellular locations. LRRs
 correspond to structural units consisting of a beta strand
 (LxxLxxN/CxL conserved pattern) and an alpha helix. This
 alignment contains 12 strands corresponding to 11 full
 repeats, consistent with the extent observed in the
 subfamily acting as Ran GTPase Activating Proteins
 (RanGAP1)"
 /db_xref="CDD:cd00116"

ORIGIN
 Query Match 82.9%; Score 2577; DB 9; Length 3507;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2817; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 289 GATCCCAGGAACTACAGGACTATGTCCGAGGAAATTTCCGGCTCATGGAAGACCGC 348
Db 489 GATCCCAGGAACTACAGGACTATGTCCGAGGAAATTTCCGGCTCATGGAAGACCGC 548
QY 349 AATCGCGCTTAGGGGAATGTCAAACCTACGACCGGTACACCGGCTCCTCTGGTG 408
Db 549 AATCGCGCTTAGGGGAATGTCAAACCTACGACCGGTACACCGGCTCCTCTGGTG 608
QY 409 AAGGAGCACTCAAAACCCCATGAGTCCAGAGTCCAGAGTCTTGGACACAGGCGGGGACAC 468
Db 609 AAGGAGCACTCAAAACCCCATGAGTCCAGAGTCCAGAGTCTTGGACACAGGCGGGGACAC 668
QY 469 GCGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTGGAGCCAGAC 528
Db 669 GCGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTGGAGCCAGAC 728
QY 529 GAGGAGCGCCCGGAGCCAGCGCACCGTGTCTATGCAAGCGCGGAGGATAGGCAAG 588
Db 729 GAGGAGCGCCCGGAGCCAGCGCACCGTGTCTATGCAAGCGCGGAGGATAGGCAAG 788
QY 589 TCCATGTGGGACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTTTCCAAAGGAGA 648
Db 789 TCCATGTGGGACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTTTCCAAAGGAGA 848
QY 649 TTTGATTATCTTCTACATCACTGAGGAGATGAACAGAGTGCACAGGATGAGC 708
Db 849 TTTGATTATCTTCTACATCACTGAGGAGATGAACAGAGTGCACAGGATGAGC 908
QY 709 ATGCAAGACCTCATCTTCAGCTGTGCGCTGAGCCAGCGCGCTCTCCAGGAGCTCATC 768
Db 909 ATGCAAGACCTCATCTTCAGCTGTGCGCTGAGCCAGCGCGCTCTCCAGGAGCTCATC 968
QY 769 CGAGTTCCCGAGCGCTCTTTTCAATCATCGAGCTTCGATGAGCTCAAGCCTTCTTTC 828
Db 969 CGAGTTCCCGAGCGCTCTTTTCAATCATCGAGCTTCGATGAGCTCAAGCCTTCTTTC 1028
QY 829 CAGGATCCTCAGGACCGCTTGTGCTCTGCTGGAGAGAAACGGCCCAAGAGCTGCTT 888
Db 1029 CAGGATCCTCAGGACCGCTTGTGCTCTGCTGGAGAGAAACGGCCCAAGAGCTGCTT 1088
QY 889 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGG 948
Db 1089 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGG 1148
QY 949 CCAACGGCTTTGAGAGCTCCACCGTCTGCTGGAGCAACCCAGGCAATGGAGATCCTG 1008
Db 1149 CCAACGGCTTTGAGAGCTCCACCGTCTGCTGGAGCAACCCAGGCAATGGAGATCCTG 1208
QY 1009 GGTCTTCTGAGGACAGAAAGGAGATATCTTCAAGTATTTCCCAATGAGAGCAG 1068
Db 1209 GGTCTTCTGAGGACAGAAAGGAGATATCTTCAAGTATTTCCCAATGAGAGCAG 1268
QY 1069 GCGGGCAAGCTTCAATTACGTGAGGAGCAAGAGCTCTTTCAACATGTGCTTCGTC 1128
Db 1269 GCGGGCAAGCTTCAATTACGTGAGGAGCAAGAGCTCTTTCAACATGTGCTTCGTC 1328
QY 1129 CCGCTGGTGTGGTGTGTGTAACCTGCTCAGAGCAGCTGAGAGGTGGGGGCTG 1188
Db 1329 CCGCTGGTGTGGTGTGTGTAACCTGCTCAGAGCAGCTGAGAGGTGGGGGCTG 1388
QY 1189 TTGAGACAGAGCTCCAGGACCACTGCAAGTGTACATGCTCTACCTGAGTCTGATG 1248
Db 1389 TTGAGACAGAGCTCCAGGACCACTGCAAGTGTACATGCTCTACCTGAGTCTGATG 1448
QY 1249 CAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCAAGAGAGGTTGTGCTCC 1308
Db 1449 CAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCAAGAGAGGTTGTGCTCC 1508
QY 1309 TTGGCGGACAGTGGGCTCTGGAATCAGAAATCCTATTTGAGGAGCAGGACCTCCGGAAG 1368
Db 1509 TTGGCGGACAGTGGGCTCTGGAATCAGAAATCCTATTTGAGGAGCAGGACCTCCGGAAG 1568
QY 1369 CAGGCGCTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1428

Db 1569 CACGCGCTAGACGGGGAAGACGTCTCTGCTCTTCTCAACATGAACATCTTCCAGAGGAC 1628
QY 1429 ATCAACTGTGAGAGTACTACAGCTTCACTTGTAGTTTCCAGGAATCTTTTCAGCT 1488
Db 1629 ATCAACTGTGAGAGTACTACAGCTTCACTTGTAGTTTCCAGGAATCTTTTCAGCT 1688
QY 1489 ATGTACTATATCTCTGGACGAGGGGAGGGGGGAGCCAGACCCAGAGAGCTGACACAGG 1548
Db 1689 ATGTACTATATCTCTGGACGAGGGGAGGGGGGAGCCAGACCCAGAGAGCTGACACAGG 1748
QY 1549 CTGTGTACCGAGTACGCGTTTCTGAAAGAGTCTTCTGGCACTTCAACAGCGCTTCTG 1608
Db 1749 CTGTGTACCGAGTACGCGTTTCTGAAAGAGTCTTCTGGCACTTCAACAGCGCTTCTG 1808
QY 1609 TTTGGACTCTGTACGAGGAGACCCAGAGCCACCTGGAGAGAGTCTCTGCTGAAAGTCT 1668
Db 1809 TTTGGACTCTGTACGAGGAGACCCAGAGCCACCTGGAGAGAGTCTCTGCTGAAAGTCT 1868
QY 1669 TCGCCGCACATCAAGATGAGACCTGTTGAGTGGATCCAAAGCAAGCTCAAGAGCAGCGC 1728
Db 1869 TCGCCGCACATCAAGATGAGACCTGTTGAGTGGATCCAAAGCAAGCTCAAGAGCAGCGC 1928
QY 1729 TCCACCTGTGAGAGGCTCCTTGGAGTCTTCAAGTGTCTTGTACGAGATCCAGAGAGAG 1788
Db 1929 TCCACCTGTGAGAGGCTCCTTGGAGTCTTCAAGTGTCTTGTACGAGATCCAGAGAGAG 1988
QY 1789 GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGGTGATCGTGTGTCAGCAACATTTGCCCTC 1848
Db 1989 GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGGTGATCGTGTGTCAGCAACATTTGCCCTC 2048
QY 1849 AAGTGGAGCACATGCTCTCTCTGCTGAAGCGCTGCAAGAGCGCGCAGGTCGTG 1908
Db 2049 AAGTGGAGCACATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2108
QY 1909 CACTTGTATGGCGCACCTACAGCGGAGACGGGGAAGACCGCGAGGTCTCCCGCAGGA 1968
Db 2109 CACTTGTATGGCGCACCTACAGCGGAGACGGGGAAGACCGCGAGGTCTCCCGCAGGA 2168
QY 1969 GCGCACACGCTGTGTGTCAGCTCAGACAGAGAGGACCGTCTGCTGAGAGCTCAGCT 2028
Db 2169 GCGCACACGCTGTGTGTCAGCTCAGACAGAGAGGACCGTCTGCTGAGAGCTCAGCT 2225
QY 2029 GAACATCTGCGAGCGGCTCTGTCACCAATCCAAACCTGATAGAGTGTCTCTGATCCGA 2088
Db 2226 GAACATCTGCGAGCGGCTCTGTCACCAATCCAAACCTGATAGAGTGTCTCTGATCCGA 2285
QY 2089 AATGCGCTGGGACCGCGGGGTGAAGTGTCTCTGTCAGAGACTCAGACACCCCACTGC 2148
Db 2286 AATGCGCTGGGACCGCGGGGTGAAGTGTCTCTGTCAGAGACTCAGACACCCCACTGC 2345
QY 2149 AAATTTCAAGAACTGAGGCTGAAGAGTGGCGCATCTCCAGCTCAGCTCGGAGGACCTC 2208
Db 2346 AAATTTCAAGAACTGAGGCTGAAGAGTGGCGCATCTCCAGCTCAGCTCGGAGGACCTC 2405
QY 2209 TCTGAGCTCTCATAGCCAAATGAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2268
Db 2406 TCTGAGCTCTCATAGCCAAATGAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2465
QY 2269 GGNATTTCCAGGACATGCTGCTTTTGGAGGGCTGGGGCATCCCGAGTCCAGGCTGACG 2328
Db 2466 GGNATTTCCAGGACATGCTGCTTTTGGAGGGCTGGGGCATCCCGAGTCCCGAGGCTGACG 2525
QY 2329 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTG 2388
Db 2526 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTG 2585
QY 2389 CTGGCAACCAACCATCTGTTGAGTTGAGTGGACCTGACAGGAATGCACCTGGAGGATTTG 2448
Db 2586 CTGGCAACCAACCATCTGTTGAGTTGAGTGGACCTGACAGGAATGCACCTGGAGGATTTG 2645
QY 2449 GGCCTCAGCTTACTATGCGAGGACTGAGGACCCAGTCTGAGACTACGAGCTTTGTGG 2508

```
Db 2646 GGCCTGAGGTTATATGCCAGGAGCTAGGACCAACCCAGTCTGCAGACTAGCGACTTTGTGG 2705
QY 2509 CTGAAGATCTGCGGCTCACTGCTGCTGCTGAGAGAGCTGGCTCAACTCTCAGTGTG 2568
Db 2706 CTGAAGATCTGCGGCTCACTGCTGCTGCTGAGAGAGCTGGCTCAACTCTCAGTGTG 2765
QY 2569 AACACAGACCTGAGAGAGCTGAGACCTGAGCCTGAATGAGCTGGGGAGCCTCGGGTGTG 2628
Db 2766 AACACAGACCTGAGAGAGCTGAGACCTGAGCCTGAATGAGCTGGGGAGCCTCGGGTGTG 2825
QY 2629 CTGCTGTGAGGGCTCAGGATCCAGCTGCAAGCTCCAGAGCTCCAGACCTCGGGTGTG 2688
Db 2826 CTGCTGTGAGGGCTCAGGATCCAGCTGCAAGCTCCAGAGCTCCAGACCTCGGGTGTG 2885
QY 2689 TGCGGCTGGGCTGCGGCTGCTGAGAGCTTCTGCTGGTGTCTCCAGGCAACACACAC 2748
Db 2886 TGCGGCTGGGCTGCGGCTGCTGAGAGCTTCTGCTGGTGTCTCCAGGCAACACACAC 2945
QY 2749 CTCCGGAGCTGAGCTTGAATTTCAACGACCTGGGAGAGCTGGGGCTGTGTTGCTGGCT 2808
Db 2946 CTCCGGAGCTGAGCTTGAATTTCAACGACCTGGGAGAGCTGGGGCTGTGTTGCTGGCT 3005
QY 2809 GAGGGCTGCAACATCCGCTGCAAGCTCCAGAACTGTGCTGGATGAGCTGTGGCTC 2868
Db 3006 GAGGGCTGCAACATCCGCTGCAAGCTCCAGAACTGTGCTGGATGAGCTGTGGCTC 3065
QY 2869 ACAGCAAGGCTGTGAGAACTTTTACTTCAACCTGGGAGATCAACGACCTTGACCGAC 2928
Db 3066 ACAGCAAGGCTGTGAGAACTTTTACTTCAACCTGGGAGATCAACGACCTTGACCGAC 3125
QY 2929 CTTTACCTGACCAACAGCCCTAGGGAGACACAGGTGTCGAGCTGCTTTGCAAGCGGCTG 2988
Db 3126 CTTTACCTGACCAACAGCCCTAGGGAGACACAGGTGTCGAGCTGCTTTGCAAGCGGCTG 3185
QY 2989 AGCCATCTGCTGCAAACTCCGAGTCTCTGCTGTTATTTGGATGAGACCTGAATAAATG 3048
Db 3186 AGCCATCTGCTGCAAACTCCGAGTCTCTGCTGTTATTTGGATGAGACCTGAATAAATG 3245
QY 3049 ACCACAGTGTGGAGCGCTTCGAGTAAACAACTTTATTTGGACATTTGGCTGCTGA 3108
Db 3246 ACCACAGTGTGGAGCGCTTCGAGTAAACAACTTTATTTGGACATTTGGCTGCTGA 3305

RESULT 5
LOCUS AV116205
DEFINITION Homo sapiens monarch-1 splice form II mRNA, linear PRI 21-MAY-2003
alternatively spliced.
ACCESSION AV116205
VERSION AV116205.1 GI:21711822
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3563)
Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.
Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat
Protein that Controls Classical and Nonclassical MHC Class I Genes
J. Immunol. 170 (11), 5354-5358 (2003)
12759408
2 (bases 1 to 3563)
Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
Direct Submission
AUTHORS
TITLE Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA
JOURNAL
PUBMED Location/Qualifiers
1. 3563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
```

```
CDS
221..3241
/contig="U937"
/feature="monarch-1 splice form II"
/protein_id="AA075143.1"
/db_xref="GI:21711823"
/translation="MLRTAGDGLCRSLTYLEEALEVKPKPLVLTATLGEKPKIP
WSEKAGPLEMAQLIIFHFGPEANRLALSTFFERINRDLWEROEDRVDTPPGG
PSSNQSTCLEVSLVTPRQDPTQYDYVRKFLMEDNRNARCEVNLSHRYRL
LLVXHSNPMQVQQLDTRGHARTVGHQASPIKTIETLFEDEPFPPTVVMQGA
AGIGKMLAHKMLMDADKLFQGRFDYLFYINCRMNQSAETCSMDLIFSCVPEPS
APLOELIRVPERLLFIIDGDELKPSFHDPOGMCLEWEEKRPTLLSLRKLPLP
ELSLIITPTALEKLHLEHPRHVEILGFSEAEKEFYKYFHNAGQAGOVNVR
DNEPLTFPCFVPLVWVCTCLOQLEGGLLRQTSRTTAYMYLILSLMQPKGAP
DLPFPNPGIUSLAADGLMNQKILFEEQDKKGGEDVSFALNMNIFQDKINCR
YYSITHSFQFFAAMYILDEGGAGPDQDVTLLTEYASERSFALSRFLFGL
LINESTRHLEKSLKMWSPHIMDLQWIKQAQSDGSLQOQSLEFSCVYEIEEE
FIQALSHFOVIVSNIAKMEHWSFCLKRCRQAOVLHLYGATYADGBDRACSA
GAHTLLVLRPERTVLLDAYSEHIAAALCTNPNIIELSLYRNALSGRGVKKLCOGLRH
PCKLONLRKRECISSACEDLSAALIANLNRMDLSGNGVFPFMMILCEGLRHP
QCRLOMQLRKQLESAGCQEMASVLTGNPHLVELDLTGNALDELGLLCOGLRHPV
CRLTLMLKI CRLTAACDELASTLSVQNSRELDLSLNELDGLVLLCGLRHPV
KLOTIRLIGICRISAAECGLSVLQANHLRELDLSFNDLGDWGLWLLAEGHQHAPCR
LQKLWLLFGMDLNTKTHSLAALRVTKPYLDIGC"
misc_feature 869..892
/feature="Region: P-loop"
misc_feature 2069..3238
/feature="Region: leucine rich repeat"
misc_feature 221..289
/feature="Region: pyrin domain"

ORIGIN
Query Match 82.4%; Score 2562; DB 9; Length 3563;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 289 GATCCCCAGGAACTACAGGAGCTATCTCGCAGGAAATTCGGCTCATGGAAGACCGC 348
DB 590 GATCCCCAGGAACTACAGGAGCTATCTCGCAGGAAATTCGGCTCATGGAAGACCGC 649
QY 349 AATGCGCCCTAGGGAAATGTCAACCTCAGCACCAGCTACACCGGCTCTCTGCTGGTG 408
DB 650 AATGCGCCCTAGGGAAATGTCAACCTCAGCACCAGCTACACCGGCTCTCTGCTGGTG 709
QY 409 AAGGAGCTCAAAACCCCATGAGGTCAGGAGCTTCTGGAACAGGCGGGGACAC 468
DB 710 AAGGAGCTCAAAACCCCATGAGGTCAGGAGCTTCTGGAACAGGCGGGGACAC 769
QY 469 GCGAGGCGGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 528
DB 770 GCGAGGCGGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 829
QY 529 GAGGAGCGCCCGAGCCACCGCGACCGCTGTGTGTCAGCAAGCGCGGAGGATAGGCAAG 588
DB 830 GAGGAGCGCCCGAGCCACCGCGACCGCTGTGTGTCAGCAAGCGCGGAGGATAGGCAAG 889
QY 589 TCCATGTGGGACACAAAGGTGATGCTGAGTGGGCGGAGCGGAACTCTTCCAAAGGAGA 648
DB 890 TCCATGTGGGACACAAAGGTGATGCTGAGTGGGCGGAGCGGAACTCTTCCAAAGGAGA 949
QY 649 TTTGATTTATCTTCTTACATCACTGAGGAGATGAACAGGATGACCGGATGCGAGC 708
DB 950 TTTGATTTATCTTCTTACATCACTGAGGAGATGAACAGGATGACCGGATGCGAGC 1009
QY 709 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGGAGCTCATC 768
DB 1010 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGGAGCTCATC 1069
QY 769 CGAGTTCGCGAGCGCTCTTTCATCATCATGAGCGCTTCGATGAGCTCAAGCTCTTCTTC 828
DB 1070 CGAGTTCGCGAGCGCTCTTTCATCATCATGAGCGCTTCGATGAGCTCAAGCTCTTCTTC 1129
```

QY 829 CACGATCCTCAGGAGCCCTGGTGGCTCTGCTGGAGAGAGAAACGGCCCAACGAGAGTGGCTT 888
Db 1130 CACGATCCTCAGGAGCCCTGGTGGCTCTGCTGGAGAGAGAAACGGCCCAACGAGAGTGGCTT 1189
QY 889 CTTAAACAGCTTAATTCGGAAGAGTGTCTCCCTGAGCTATCTTTGCTCATCACCACACGG 948
Db 1190 CTTAAACAGCTTAATTCGGAAGAGTGTCTCCCTGAGCTATCTTTGCTCATCACCACACGG 1249
QY 949 CCCACGGCTTTGGAGAGCTCCACCGTCTCTGCTGGAGACACCCAGGCAATGTGGAGATCCTG 1008
Db 1250 CCCACGGCTTTGGAGAGCTCCACCGTCTCTGCTGGAGACACCCAGGCAATGTGGAGATCCTG 1309
QY 1009 GGCTTCTCTGAGGAGAGAGAGAGATATCTTCAAGATATTTCCAAATGACAGAGCAG 1068
Db 1310 GGCTTCTCTGAGGAGAGAGAGAGATATCTTCAAGATATTTCCAAATGACAGAGCAG 1369
QY 1069 GCGGGCCAACTCTCAATACGTGAGGAGCAACGAGGCTCTTTCACCATGTCTTCGTC 1128
Db 1370 GCGGGCCAACTCTCAATACGTGAGGAGCAACGAGGCTCTTTCACCATGTCTTCGTC 1429
QY 1129 CCCTGGTGTGTGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1188
Db 1430 CCCTGGTGTGTGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1489
QY 1189 TTGAGACAGAGCTCCAGGACCAACCTGCACTGTGATGATGCTCTACCTGCTGAGTCTGATG 1248
Db 1490 TTGAGACAGAGCTCCAGGACCAACCTGCACTGTGATGATGCTCTACCTGCTGAGTCTGATG 1549
QY 1249 CAACCCAAAGCCGGGGCCCGCGCTCCAGCCGCCACCCCAACAGAGAGAGGTTGTGCTCC 1308
Db 1550 CAACCCAAAGCCGGGGCCCGCGCTCCAGCCGCCACCCCAACAGAGAGAGGTTGTGCTCC 1609
QY 1309 TTGGCGCAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 1368
Db 1610 TTGGCGCAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 1669
QY 1369 CAGGCTTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTCCAGAGAGAC 1428
Db 1670 CAGGCTTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTCCAGAGAGAC 1729
QY 1429 ATCAACTGTGAGAGTACTACAGCTTCACTCCACTTGAGTTTCCAGGAATCTTTGACAGCT 1488
Db 1730 ATCAACTGTGAGAGTACTACAGCTTCACTCCACTTGAGTTTCCAGGAATCTTTGACAGCT 1789
QY 1489 ATGTACTATATCTTGGACGAGGGGAGGGGCGGAGCCAGACCCAGAGCTGACCAAG 1548
Db 1790 ATGTACTATATCTTGGACGAGGGGAGGGGCGGAGCCAGACCCAGAGCTGACCAAG 1849
QY 1549 CTGTTGACGAGTACGCTTTCTGAAAGAGAGCTCTCTGGCACTCAGCCGCTTCTG 1608
Db 1850 CTGTTGACGAGTACGCTTTCTGAAAGAGAGCTCTCTGGCACTCAGCCGCTTCTG 1909
QY 1609 TTTGGACTCTGAAACGAGGAGACCAAGGACCACTGAGAGAGTCTCTGCTGGAAGTCT 1668
Db 1910 TTTGGACTCTGAAACGAGGAGACCAAGGACCACTGAGAGAGTCTCTGCTGGAAGTCT 1969
QY 1669 TCGCCGACATCAAGATGAGACCTTCTGAGTGGATCCAAAGCAAAAGCTCAGAGCAGCGC 1728
Db 1970 TCGCCGACATCAAGATGAGACCTTCTGAGTGGATCCAAAGCAAAAGCTCAGAGCAGCGC 2029
QY 1729 TCCACCTGACAGAGGCTCTCTGGAGTCTTCACTGCTGCTTACAGATCCAGGAGAG 1788
Db 2030 TCCACCTGACAGAGGCTCTCTGGAGTCTTCACTGCTGCTTACAGATCCAGGAGAG 2089
QY 1789 GAGTTATCCAGAGGCTCTGAGCACTTCCAGGTGATCTGGTGCAGCAATTCCTCTCC 1848
Db 2090 GAGTTATCCAGAGGCTCTGAGCACTTCCAGGTGATCTGGTGCAGCAATTCCTCTCC 2149
QY 1849 AAGATGAGCAGATGCTCTCTGCTGCTGAGGCTGTCAGAGAGGCGCCAGGCTGCTG 1908
Db 2150 AAGATGAGCAGATGCTCTCTGCTGCTGAGGCTGTCAGAGAGGCGCCAGGCTGCTG 2209
QY 1909 CACTTGTATGGCCCACTTACAGCGGAGCGGAGAGACCGCGGAGTGTCTCCGAGGA 1968

Db 2210 CACTTGTATGGCCCACTTACAGCGGAGCGGGAAGACCGCGAGGTCTCCGAGGA 2269
QY 1969 GCGCACACAGCTGTGTGGTGCAGCTCAGACAGAGAGAGACCGCTTCTGTGAGCGCTACAGT 2028
Db 2270 GCGCACACAGCTGTGTGGTGCAGCTCAGACAGAGAGAGACCGCTTCTGTGAGCGCTACAGT 2329
QY 2029 GAACATCTGCGACGCGGCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTGTGACGA 2088
Db 2330 GAACATCTGCGACGCGGCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTGTGACGA 2389
QY 2089 AATGCCCTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2148
Db 2390 AATGCCCTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2449
QY 2149 AAACCTTCAGAACCTGAGGCTGAAGAGGTGCGGATCTCCAGCTCAGCTCGAGGACCTC 2208
Db 2450 AAACCTTCAGAACCTGAGGCTGAAGAGGTGCGGATCTCCAGCTCAGCTCGAGGACCTC 2509
QY 2209 TCTGAGCTCTCATAGCAATAAGATTTGCAAGGATGATCTCAGTGGCAACCGGCTT 2268
Db 2510 TCTGAGCTCTCATAGCAATAAGATTTGCAAGGATGATCTCAGTGGCAACCGGCTT 2569
QY 2269 GGAATCCCGAGCATGATGCTGCTTGGAGGGCTCGGCGATCCCGAGTGCAGCTCGAG 2328
Db 2570 GGAATCCCGAGCATGATGCTGCTTGGAGGGCTCGGCGATCCCGAGTGCAGCTCGAG 2629
QY 2329 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTG 2388
Db 2630 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTG 2689
QY 2389 CTGCGCAACCAACCAATCTGGTTGAGTGGACCTGACAGAAATGCACTGGAGATTTG 2448
Db 2690 CTGCGCAACCAACCAATCTGGTTGAGTGGACCTGACAGAAATGCACTGGAGATTTG 2749
QY 2449 GGCCTGAGGTTACTATGCCAGGACTGAGGACCCAGTCTGACAGTACGAGCTTTGTGG 2508
Db 2750 GGCCTGAGGTTACTATGCCAGGACTGAGGACCCAGTCTGACAGTACGAGCTTTGTGG 2809
QY 2509 CTGAAGATCTGCCGCTCTCAGTGTGCTGTGACAGCTGGCTCAACTCTCAGTGTG 2568
Db 2810 CTGAAGATCTGCCGCTCTCAGTGTGCTGTGACAGCTGGCTCAACTCTCAGTGTG 2869
QY 2569 AACAGAGCTTGAGAGAGCTGGACCTGAGCTGAAATGAGCTGGGGGACCTCGGGGTGCTG 2628
Db 2870 AACAGAGCTTGAGAGAGCTGGACCTGAGCTGAAATGAGCTGGGGGACCTCGGGGTGCTG 2929
QY 2629 CTGCTGTGAGGGCTCAGGCACTCCACGTGCAAGCTCCAGACCTCGCGTGGGATC 2688
Db 2930 CTGCTGTGAGGGCTCAGGCACTCCACGTGCAAGCTCCAGACCTCGCGTGGGATC 2989
QY 2689 TGCGGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTGTGCTTCCAGGGCAACCAAC 2748
Db 2990 TGCGGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTGTGCTTCCAGGGCAACCAAC 3049
QY 2749 CTCGGAGCTGAGCTTGAAGTTCAACGACTGGGAGACTGGGGCTGTGGTCTGGCT 2808
Db 3050 CTCGGAGCTGAGCTTGAAGTTCAACGACTGGGAGACTGGGGCTGTGGTCTGGCT 3109
QY 2809 GAGGGCTGCAACATCCCGCTGCAGACTCCAGAACTGTGG 2850
Db 3110 GAGGGCTGCAACATCCCGCTGCAGACTCCAGAACTGTGG 3151

RESULT 6
AY095146
LOCUS AY095146
DEFINITION Homo sapiens PYRIN-containing APAF1-like protein 7 mRNA, complete cds.
ACCESSION AY095146
VERSION AY095146.1 GI:21314906
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Wang, L., Manji, G.A., Grenier, J.M., Al-Garawi, A., Merriam, S., Lora, J.M., Geddes, B.J., Briskin, M., Distefano, P.S. and Bertin, J.
TITLE XPAPAF7, a Novel PYRIN-containing Apaf1-like Protein That Regulates Activation of NF-kappa B and Caspase-1-dependent Cytokine Processing
JOURNAL J. Biol. Chem. 277 (33), 29874-29880 (2002)
PUBMED 12019269
REFERENCE 2 (bases 1 to 3827)
AUTHORS Bertin, J.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2002) Inflammation, Millennium Pharmaceuticals Inc., 45 Sidney Street, Cambridge, MA 02139, USA
FEATURES Location/Qualifiers
source
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
177..3382
/notes="PYPAF7; PYRIN-containing nucleotide-binding site/leucine-rich repeat family member"
/codon_start=1
/product="PYRIN-containing APAF1-like protein 7"
/protein_id="AAM18227.1"
/db_xref="GI:21314907"
/translations="MURTAGRDGLCRLSYVLEEAVALKKPKLYLGATLGEKGKIPWSMEXAGLEMAQLLTHFGPEAWRLALSTFERINRKLWERGQRELDVDPGGPSLGNQSTCLLEVSIVTKPQDETQYVRKRLMEDNRNARLCEVCNLSHRVTRLPLVKSHSNPQVQQLDTRGHARTVGHQASPKITLFPDEPERPEPTVVMGQAGTIGKSLAHKWLMDAGLKFGREDYLFYINCRMNQSAQTECSQMDLIFSCPEPSAPLOELIIVPERLLFIDGDELPKSPHDQFQPCWLCKWEKPTLNLNLSLRKLLPLELLITPTALEKHLRHPHVEIILGFEAERKEYFYFYFHNAAQGGQVFNVRLQPNPDMFCVPLAVWCCTQQLQEGGLLRQTSRTTAVMYLLSLMQPKGAPRIQPTNPGRGSLAADGLWNQKILPEEQDLRKHGDGDSVAFILNMNIFQKDINCRYFSFHLSPFERPAAAMYILDEGGGAGPDQVTRLLTEVAFSEFSFLATLSRFLFGLLINEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQSDGSLTQOGLSEFFSCLVEIQBEFTQQAUSHQVIVSNIAKSMHWSSFCIKKCRSAQVHLHYGATYSADGEDRARCASGAHTLLQIPERTVLDDAYSEHLAAALCTNPMLISLRYNALSGSGVLLCCQGRHPNCKLQWLRIKRLISSACDLSAALIANKNLTLDLSNGVGFPGMVLCCGLRHQPCLQMTQLRKQLESQACEMASVLGTNPHLVELDLTGNALDGLRLILCCGLRHVPCLRLTWLTKRLTAAACDEBLASTVSNLSRLVELDLSNELGDLVLLICEGLRHPTKCLQTLRLGICRLGSAACEGSLVULOANHLRELDLSFNDIGDMGLVLLASGLQHPACRLQKLWLDSCGLITAKACENLFTFGINQTLITDLYLTNNALGDTGVRLLCRKLSPHGCKILRVLWFGMDLKNMTHSRLLAALRVTKPYLDIGC"
ORIGIN
Query Match 79.6%; Score 24.75; DB 9; Length 3827;
Best Local Similarity 9.8%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 289 GATCCCAGGAAACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 348
Db 546 GATCCCAGGAAACCTACAGGACTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 605
QY 349 AATCGCGGCTAGGGGAAATGTGTCAACTCAGCCACGGGTACACCCGGCTCTCTGTGGTG 408
Db 606 AATGCGCGCTCAGGGGAAATGTGTCAACTCAGCCACGGGTACACCCGGCTCTCTGTGGTG 665
QY 409 AAGGAGCACTCAAAACCCATGCAAGTCCAGCAGCAGCTTCTTGACACAGCGCGGACAC 468
Db 666 AAGGAGCACTCAAAACCCATGCAAGTCCAGCAGCAGCTTCTTGACACAGCGCGGACAC 725
QY 469 GCAGGAGACCGTGGGACACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 528
Db 726 GCAGGAGACCGTGGGACACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 785
QY 529 GAGGAGCGCCCGGAGCCACCGCGCACCGTGGTTCATGCAAGCGCGGAGGATAGCAAG 588

1669 TCCTCCGACATCAAGATGAGACCTGTTGACGTGGATCAAAAGCAAAAGCTCAGAGCGAGCGC 1728
1926 TCCTCCGACATCAAGATGAGACCTGTTGACGTGGATCAAAAGCAAAAGCTCAGAGCGAGCGC 1985
1729 TCACACCTGACAGAGGCTCTTGGAGTCTTCTGAGTCTGTTGACGAGATCCAGGAGGAG 1788
1986 TCACACCTGACAGAGGCTCTTGGAGTCTTCTGAGTCTGTTGACGAGATCCAGGAGGAG 2045
1789 GAGTTTATCCAGAGGCTCTGAGCCTTCCAGTGTATCTGGTGCAGCAATTCGCTCC 1848
2046 GAGTTTATCCAGAGGCTCTGAGCCTTCCAGTGTATCTGGTGCAGCAATTCGCTCC 2105
1849 AAGATGAGACATGAGTCTCTCTGTTCTGTTGAAAGCGCTGACAGGAGCGCCAGGTCGT 1908
2106 AAGATGAGACATGAGTCTCTCTGTTCTGTTGAAAGCGCTGACAGGAGCGCCAGGTCGT 2165
1909 CACTTGTATCCGCGCCACTACAGCGGAGCGGAGAGAGCGCGGAGGTCGTCGAGGA 1968
2166 CACTTGTATCCGCGCCACTACAGCGGAGCGGAGAGAGCGCGGAGGTCGTCGAGGA 2225
1969 GCGCACACGCTGTTGTTGACGCTCAGACACAGAGAGAGCGGTCGTCGAGCGCTACAGT 2028
2226 GCGCACACGCTGTTGTTGACGCT---ACGAGAGAGACCGTTCGCTGAGCGCTACAGT 2282
2029 GACATCTGGAGCGGCTCTGTGACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2088
2283 GAACATCTGGAGCGGCTCTGTGACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2342
2089 AATGCCCTGGAGCGGCTCTGTGACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2148
2343 AATGCCCTGGAGCGGCTCTGTGACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2402
2149 AAACCTTCAGAACCTGAGCTGAAAGAGTGGCGCATCTCCAGCTCAGCGCTCGAGGACCTC 2208
2403 AAACCTTCAGAACCTGAGCTGAAAGAGTGGCGCATCTCCAGCTCAGCGCTCGAGGACCTC 2462
2209 TCTGAGCTCTATAGCAATAGAAATTTGACAGGATGATCTCAGTGGCAACGGCGTT 2268
2463 TCTGAGCTCTATAGCAATAGAAATTTGACAGGATGATCTCAGTGGCAACGGCGTT 2522
2269 GGATTCACGAGCATGATGCTGCTTGGAGGCTGGCGCATCCCACTCAGCTCGAG 2328
2523 GGATTCACGAGCATGATGCTGCTTGGAGGCTGGCGCATCCCACTCAGCTCGAG 2582
2329 ATGATTCAGTTGAGGAGTCTCAGCTGAGCTGCGGCTTGTGACGAGATGCTCTGTG 2388
2583 ATGATTCAGTTGAGGAGTCTCAGCTGAGCTGCGGCTTGTGACGAGATGCTCTGTG 2642
2389 CTGGCACCAACCAATCTGTTGAGTGGACCTGACGAGAAATGCACTGGAGGATTTG 2448
2643 CTGGCACCAACCAATCTGTTGAGTGGACCTGACGAGAAATGCACTGGAGGATTTG 2702
2449 GGCTCAGGTTACTATGCCAGGAGCTGAGCAGCAGCTGACAGCTACCGGACTTTCTGG 2508
2703 GGCTCAGGTTACTATGCCAGGAGCTGAGCAGCAGCTGACAGCTACCGGACTTTCTGG 2762
2509 CTGAAGATCTGCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
2763 CTGAAGATCTGCCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2822
2569 AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2628
2823 AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2882
2629 CTGCTGTGAGGAGCTCAGGACATCCCACTGCAAGCTCCAGACCTCGCGGTTGGGATC 2688
2883 CTGCTGTGAGGAGCTCAGGACATCCCACTGCAAGCTCCAGACCTCGCGGTTGGGATC 2942
2689 TGGCGGTGGGCTCTGCGGCTGTGAGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2748
2943 TGGCGGTGGGCTCTGCGGCTGTGAGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3002

2749 CTCGGAGCTGACTGACTTTCAACGACCTGGAGACTGGGCTGTGTTGCTGCTGCT 2808
3003 CTCGGAGCTGACTGACTTTCAACGACCTGGAGACTGGGCTGTGTTGCTGCTGCT 3062
2809 GAGGGCTGCAACATCCGCTGAGACTCCAGAACTGTTGGCTGATAGCTGTTGGCTC 2868
3063 GAGGGCTGCAACATCCGCTGAGACTCCAGAACTGTTGGCTGATAGCTGTTGGCTC 3122
2869 ACAGCCAAAGCTTGTGAGAACTTTTACTTCCCTGGGATCAACAGACCTTGACCGAC 2928
3123 ACAGCCAAAGCTTGTGAGAACTTTTACTTCCCTGGGATCAACAGACCTTGACCGAC 3182
2929 CTTTACTGACCAACAGCGCTTAGGGGACACAGGTGTCGACTGCTTTGCAAGCGCTG 2988
3183 CTTTACTGACCAACAGCGCTTAGGGGACACAGGTGTCGACTGCTTTGCAAGCGCTG 3242
2989 AGCATCTGCTGCAAACTCCGAGTCTCTGTTTATTTGGATGACCTGATATAAATG 3048
3243 AGCATCTGCTGCAAACTCCGAGTCTCTGTTTATTTGGATGACCTGATATAAATG 3302
3049 ACCACAGTAGTGGCAGCGCTTCGAGTAACAAACCTTATTTGGACATTTGGCTGCTGA 3108
3303 ACCACAGTAGTGGCAGCGCTTCGAGTAACAAACCTTATTTGGACATTTGGCTGCTGA 3362

RESULT 7
AY154467
LOCUS 3827 bp mRNA linear PRI 20-FEB-2003
DEFINITION Homo sapiens NALP12 (NALP12) mRNA, complete cds.
ACCESSION AY154467
VERSION AY154467.1 GI:28436377
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3827)
AUTHORS Tschoopp,J., Martinon,F. and Burns,K.
TITLE NALPs: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
MEDLINE 22451042
PUBMED 12563287
REFERENCE 2 (bases 1 to 3827)
AUTHORS Martinon,F., Hofmann,K. and Tschoopp,J.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
source
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3827
/gene="NALP12"
177..3362
/gene="NALP12"
/notes="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP12"
/protein_id="AA018163.1"
/db_xref="GI:28436378"
/translation="MLRTAGRDGLRLSTYLEEAEVLEKKFLKLYLGTATELGEGKIP
PSSMGKAGPLEMAQLITHTFPEAEWRALSTFERINRDLWEGQREDIVRTPPGG
LLVKEHSDMQVQQLDTRGHARTVGHQASPIKLTLEPDEERPEPTVVMQGA
AGTGSMLAHKMLMDWADKLFQGRDYLFYINCRMNQSAQSCQDILFSCWPEPS
APQLGIRFVERLLFTIDGDELPKSFHDQFQWCLCWEERPTLELNLIRKKLHP
EUSLLITRTRALEKHLRHLRHEIILGFSABRKEYFYKVFHNAEQAGQVFNVR
DNEPLFTMCVPVLCVWVCTCLQQLEGLLRQRTSTTTAVVLYLLSLQMPKPGAP
RLQPPNQRGLCSLAADGLWNKILFEEDQRLKHLGDGSDVAFNNMIFQKDNCR
YYSFIHLSFOEPFAAMYIILDEGGAGPDQVTRLITLTVAFSERSLATLSRFLPGL
LNESTRSHLEKSLCMKVSPIKMDLLQWIKSQASDGSSTLQQGSLEFFSCLVEIQEE

FIOALXSHPOIVVSNIAKSHWYSSFCFLKCSAOLVHLYGATVYSDAGDEBARCSA
 GATTLAVOLPEITVSKVDEHSEALATCNPNLTELJYNALSGFVGLHCOGLRHVC
 NCKLOILKRCRISISSACCEMAYLHANKHLMVDJSGNGVSGPMMWLLCOGLRHVC
 CKXLOILKRCOLEGSAACEMAYLHANKHLMVDJSGNGVSGPMMWLLCOGLRHVC
 FLATLWIKTICRLGSAACEBLATSVNLSHRELDLSNELGDLGLGLHCOGLRHVC
 LOTLRJICRLGSAACEBLATSVNLSHRELDLSNFLDGLWGLLWGLAQHPACRL
 QKWLSDKCSACNLYETLJINOTLDTLAYTNNALGDTGVLLCKRLSHSGCKLR
 VWLFGMDLKNWTHSALATLRYTKPGLDLC"

ORIGIN

Query Match 79.6%; Score 2475; DB 9; Length 3827;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 2; Indels 3

289	QY	GATCCCGAGAAACCTTACAGGACTATGTCGCGAGGAAATTCGCGCTCATGGAGACCCG	348
546	Db	GATCCCGAGAAACCTTACAGGACTATGTCGCGAGGAAATTCGCGCTCATGGAGACCCG	605
349	QY	AATGCGCGCTAGGGGAATGTGTCAACCTCAGGCACCGGTACACCCGGCTCCTGCTGGTG	408
606	Db	AATGCGCGCTAGGGGAATGTGTCAACCTCAGGCACCGGTACACCCGGCTCCTGCTGGTG	665
409	QY	AAGGAGACTCAAACCCCATGAGGTCCAGCAGGAGCTTCTGGACACAGGCCGGGAGAC	468
666	Db	AAGGAGACTCAAACCCCATGAGGTCCAGCAGGAGCTTCTGGACACAGGCCGGGAGAC	725
469	QY	GCGAGGACCGTGGGACACCAAGGCTAGGCCCATCAAGATAGAGACCCCTCTTTGAGCCAG	528
726	Db	GCGAGGACCGTGGGACACCAAGGCTAGGCCCATCAAGATAGAGACCCCTCTTTGAGCCAG	785
529	QY	GAGGAGCGCCCGAGGCACCGCGCACCGTGGTCAATGCAAGCGCGGACAGGATAGGCAAG	588
786	Db	GAGGAGCGCCCGAGGCACCGCGCACCGTGGTCAATGCAAGCGCGGACAGGATAGGCAAG	845
589	QY	TCCATGCTGCACACAAAGGTGATGCTGGCTGGCGGACGGGAGCTCTTCCAGGCAGA	648
846	Db	TCCATGCTGCACACAAAGGTGATGCTGGCTGGCGGACGGGAGCTCTTCCAGGCAGA	905
649	QY	TTTGATTATCTCTTTACATCAACTGCAGGAGATGAACAGAGTGCACCGGAATGCAGC	708
906	Db	TTTGATTATCTCTTTACATCAACTGCAGGAGATGAACAGAGTGCACCGGAATGCAGC	965
709	QY	ATGCAAGACCTCATCTTTCAGCTGTGCTGAGCCACGCGGCGCTCTCCAGGAGCTCATC	768
966	Db	ATGCAAGACCTCATCTTTCAGCTGTGCTGAGCCACGCGGCGCTCTCCAGGAGCTCATC	1025
769	QY	CGAGTTCCCGAGCGCCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC	828
1026	Db	CGAGTTCCCGAGCGCCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC	1085
829	QY	CACGATCCTCAGGACCCCTGGTGCTGTGTGGAGAGAAACGGCCACCGGAGCTGCTT	888
1086	Db	CACGATCCTCAGGACCCCTGGTGCTGTGTGGAGAGAAACGGCCACCGGAGCTGCTT	1145
889	QY	CTTTAAGAGCTTAATTCGGAGAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG	948
1146	Db	CTTTAAGAGCTTAATTCGGAGAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG	1205
949	QY	CCCAAGCCTTTGGAGAGCTCCACCGTCTGTGTGAGCACCCCGGAGCTGTGGAGATCCTG	1008
1206	Db	CCCAAGCCTTTGGAGAGCTCCACCGTCTGTGTGAGCACCCCGGAGCTGTGGAGATCCTG	1265
1009	QY	GGCTTCTCTGAGGCAGAAAGGAAGGAATATCTTCTACAAGTATTTTCCAAATGACAGACG	1068
1266	Db	GGCTTCTCTGAGGCAGAAAGGAAGGAATATCTTCTACAAGTATTTTCCAAATGACAGACG	1325
1069	QY	GCGGGCCAAGTCTTCAATTACGTGAGGACAAACAGAGCTCTCTTTCACCATGTGCTTCGTC	1128
1326	Db	GCGGGCCAAGTCTTCAATTACGTGAGGACAAACAGAGCTCTCTTTCACCATGTGCTTCGTC	1385
1129	QY	CCCTGTGTGCTGGGTGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG	1188
1386	Db	CCCTGTGTGCTGGGTGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG	1445

1189	Qy	TTGAGACAGACGTCCTCAGGACCAACAATGCAAGTGTATACGCTCTACCTGCTGAGTCTGATG	12448
1446	Db	TTGAGACAGACGTCCTCAGGACCAACAATGCAAGTGTATACGCTCTACCTGCTGAGTCTGATG	15055
1249	Qy	CAACCCAAAGCCGGGGCCCGCGCTCCAGCCCCCAACCAAGAGAGGTTGCTGCTCC	13108
1506	Db	CAACCCAAAGCCGGGGCCCGCGCTCCAGCCCCCAACCAAGAGAGGTTGCTGCTCC	15655
1309	Qy	TTGGCGGCAGATGGGCTCTGGAATCAGAAAATCCTATTGAGGAGCAGGACCTCCCGAAG	13688
1566	Db	TTGGCGGCAGATGGGCTCTGGAATCAGAAAATCCTATTGAGGAGCAGGACCTCCCGAAG	16225
1369	Qy	CAGGCTTACAGCGGGGAAGACGCTCTGCGCTTCTCAACATGAAATCTTCCAGAAAGGAC	14248
1626	Db	CAGGCTTACAGCGGGGAAGACGCTCTGCGCTTCTCAACATGAAATCTTCCAGAAAGGAC	16895
1429	Qy	ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTTGACGCT	14888
1686	Db	ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTTGACGCT	17455
1489	Qy	ATGTACTATATCTTGACAGGGGGAGCGGGGAGCGCCACAGACGAGACGTGACACAGG	15448
1746	Db	ATGTACTATATCTTGACAGGGGGAGCGGGGAGCGCCACAGACGAGACGTGACACAGG	18005
1549	Qy	CTGTTGACGAGTACGGTTTCTGAAAGAGGCTTCTGTGCACTCACAGCGGCTTCTGCT	16088
1806	Db	CTGTTGACGAGTACGGTTTCTGAAAGAGGCTTCTGTGCACTCACAGCGGCTTCTGCT	18655
1609	Qy	TTTGGACTCTGACAGGAGACCAAGAGCCACTCGAGAGAGTCTCTCTCGGAAGGTC	16668
1866	Db	TTTGGACTCTGACAGGAGACCAAGAGCCACTCGAGAGAGTCTCTCTCGGAAGGTC	19225
1669	Qy	TCGCGGCACATCAAGATGCACTGTTGCACTGGATCCAAAGCAAAGCTCAGAGCGACGGC	17288
1926	Db	TCGCGGCACATCAAGATGCACTGTTGCACTGGATCCAAAGCAAAGCTCAGAGCGACGGC	19895
1729	Qy	TCCACCTTGACGAGGCTCTCTGGAGTCTTTCAGCTGCTTTGACGAGATCCAGGAGGAG	17888
1986	Db	TCCACCTTGACGAGGCTCTCTGGAGTCTTTCAGCTGCTTTGACGAGATCCAGGAGGAG	20445
1789	Qy	GAGTTTATCAGAGGCTCTGAGCCATTCCAGTGTATGCTGCTGAGCAATTTGCTCTCC	18448
2046	Db	GAGTTTATCAGAGGCTCTGAGCCATTCCAGTGTATGCTGCTGAGCAATTTGCTCTCC	21005
1849	Qy	AAGATGAGCACATGTCTCTCGTCTGTCTGAAGCGCTCGAGAGCGCCAGGTGCTG	19088
2106	Db	AAGATGAGCACATGTCTCTCGTCTGTCTGAAGCGCTCGAGAGCGCCAGGTGCTG	21645
1909	Qy	CACTTGTATGGCGCACCTACAGCGCGGAGCGGGGAAGACCGCGGAGGTGCTCGCAGGA	19688
2166	Db	CACTTGTATGGCGCACCTACAGCGCGGAGCGGGGAAGACCGCGGAGGTGCTCGCAGGA	22255
1969	Qy	GCGCACACGCTGTGTTGTCAGCTCAGACCAAGAGAGGACCGTTCTGCTGAGGCGCTACAGT	20288
2226	Db	GCGCACACGCTGTGTTGTCAGCTCAGACCAAGAGAGGACCGTTCTGCTGAGGCGCTACAGT	22825
2029	Qy	GAAATCTGACAGCGGCTCTGTGCAACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA	20888
2283	Db	GAAATCTGACAGCGGCTCTGTGCAACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA	23442
2089	Qy	AATGCCCTTGGCGACCGGGGGGTGAAGCTGTCTGTCAAGGATCTCAGACACCCCACTGC	21448
2343	Db	AATGCCCTTGGCGACCGGGGGGTGAAGCTGTCTGTCAAGGATCTCAGACACCCCACTGC	24002
2149	Qy	AAACTTTCAGAACTTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTTCGAGGACCTC	22088
2403	Db	AAACTTTCAGAACTTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTTCGAGGACCTC	24642
2209	Qy	TCTGAGCTCTCATAGGCAATAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGCTT	22688
2463	Db	TCTGAGCTCTCATAGGCAATAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGCTT	25222

QY 2269 GGATCCAGGCGATGCTGCTTTCGAGGGCTCGGGATCCCAAGTCAGGCTCCAG 2328
Db |||||||
QY 2523 GGATCCAGGCGATGCTGCTTTCGAGGGCTCGGGATCCCAAGTCAGGCTCCAG 2582
Db |||||||
QY 2329 ATGATTCAGTTGAGGAGTCTGAGTGGATCCGGGCTTGTGAGGAGATGGCTTCGTG 2388
Db |||||||
QY 2583 ATGATTCAGTTGAGGAGTCTGAGTGGATCCGGGCTTGTGAGGAGATGGCTTCGTG 2642
Db |||||||
QY 2389 CTGGGCAACCAACCAATCTGGTTGAGTTGGAATCTGACAGGAATGCACTGGAGGATTTG 2448
Db |||||||
QY 2643 CTTGGGCAACCAACCAATCTGGTTGAGTTGGAATCTGACAGGAATGCACTGGAGGATTTG 2702
Db |||||||
QY 2449 GGCTGAGGTACTATGCGAGGACTGAGGACCCAGTCTGACAGTACCGACTTTGTGG 2508
Db |||||||
QY 2703 GGCTGAGGTACTATGCGAGGACTGAGGACCCAGTCTGACAGTACCGACTTTGTGG 2762
Db |||||||
QY 2509 CTGAAGATCTGCGGCTCACTGCTGCTGCTGACGAGTGGCTCAACTCTCAGGTG 2568
Db |||||||
QY 2763 CTGAAGATCTGCGGCTCACTGCTGCTGCTGACGAGTGGCTCAACTCTCAGGTG 2822
Db |||||||
QY 2569 AACAGAGCTGAGAGCTGAGCTGAGCTGAATGAGTGGGGGACCTCGGGGTCTG 2628
Db |||||||
QY 2823 AACAGAGCTGAGAGCTGAGCTGAGCTGAATGAGTGGGGGACCTCGGGGTCTG 2882
Db |||||||
QY 2629 CTGCTGTGAGGCTCAGGATCCCACTGCAAGCTCAGACCTGCGGTTGGGATC 2688
Db |||||||
QY 2883 CTGCTGTGAGGCTCAGGATCCCACTGCAAGCTCAGACCTGCGGTTGGGATC 2942
Db |||||||
QY 2689 TGGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTCTGTGCTGCTCAGAGGCAACCAAC 2748
Db |||||||
QY 2943 TGGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTCTGTGCTGCTCAGAGGCAACCAAC 3002
Db |||||||
QY 2749 CTGCGGAGCTGAGTGTGAGTTCAACAGCTGAGGACTGGGGCTGTGTTGCTGGCT 2808
Db |||||||
QY 3003 CTGCGGAGCTGAGTGTGAGTTCAACAGCTGAGGACTGGGGCTGTGTTGCTGGCT 3062
Db |||||||
QY 2809 GAGGGCTGCAATCCCGCTGAGACTCCAGAACTGGCTGAGTGTGAGTGTGAGTGTG 2868
Db |||||||
QY 3063 GAGGGCTGCAATCCCGCTGAGACTCCAGAACTGGCTGAGTGTGAGTGTGAGTGTG 3122
Db |||||||
QY 2869 ACAGCAAGCTTGTGAGAACTTTTACTTCCACCTGGGATCAACAGACTTTGACCGAC 2928
Db |||||||
QY 3123 ACAGCAAGCTTGTGAGAACTTTTACTTCCACCTGGGATCAACAGACTTTGACCGAC 3182
Db |||||||
QY 2929 CTTTACTGACCAACAGCCCTAGGGGACACAGGTGTCGAGTGTGCTTGTGCAAGCGGCTG 2988
Db |||||||
QY 3183 CTTTACTGACCAACAGCCCTAGGGGACACAGGTGTCGAGTGTGCTTGTGCAAGCGGCTG 3242
Db |||||||
QY 2989 AGCCATCTGGCTGCAAACTCCGAGTCTCTGTTTATTTGGGATGACCTGAATAAATG 3048
Db |||||||
QY 3243 AGCCATCTGGCTGCAAACTCCGAGTCTCTGTTTATTTGGGATGACCTGAATAAATG 3302
Db |||||||
QY 3049 ACCACAGTAGGTGGGAGGCTTCCAGTAACAAACCTTATTGACATTGGCTGTGA 3108
Db |||||||
QY 3303 ACCACAGTAGGTGGGAGGCTTCCAGTAACAAACCTTATTGACATTGGCTGTGA 3362
Db |||||||

RESULT 8
LOCUS AY116206
DEFINITION Homo sapiens monarch-1 splice form III mRNA, complete cds;
ACCESSION AY116206
VERSION AY116206.1 GI:21711824
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3395)
AUTHORS Williams, K.L., Linhoff, M.W., Reed, W. and Ting, J.P.Y.
TITLE Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes

J. Immunol. 170 (11), 5354-5358 (2003)
12759408
REFERENCE 2 (bases 1 to 3395)
AUTHORS Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA
Location/Qualifiers
FEATURES
source

1. .3395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221. .3073

/notes="contains alternatively spliced exons 7 and 8"
/product="monarch-1 splice form III"
/protein_id="AA075144.1"
/db_xref="GI:21711825"

/translation="MLRTAGRDGLRLSTYLELEAVELEKFKLKLGLTATLGEKGKIP
WGSERAGPLLEMAQLLITHTFPEAWRLALSTFERINRDLWEGOREDLVRDTPPG
PSSLGNSCTLLEVSATVPKQDYRYVYRKRFLMEDRNARLGEVNLNRYFL
LIVKESNPMQVQQLDTRGHARTVGHQSEPKITLTFEPDEERPEPTVVMQGA
AGIGKSMIAHKVMDWADGKLFQGRFDYLYFVYINCRNOSATECSMDLIFSVMQPS
APLELIRVPERLLFIIDGDELKPFHDPQGWCLWBEKREKTELLNLIRKLLIP
ELSLITTRPTALEKLRLLEHPRHVEILGFSAEKREKVFYVFNAAQAGQVFNVR
DNEPLTMCVPLVCWVCTCLOQLEGGGLLQTSRTTAVVLYLILSLMQPKGAP
RLQPPNQRGLSLAADGLWNLKILFEBQDLRKHGLDGEDVSAFLNNIFQKIDNCR
YYSFHLSPQEFFAAMYIILDEGGAGDPDVRLLTFAYSERFLNLTSLRFLGL
LNEBTSHEKLSLKWVSPHIMKMDLQWTSQKASQDSTLQOGLSLEFSCLYIQEE
FTQALSHFQVIVSNIAISMEHWSFCLKRSQAQVHLHYGATYSADGEDARCSA
GAHTLVQLPRTVLDAVSEHLAALCTNPNLIELSLVRLNALSRSRQVLLCQGLRH
PNCKLNLRLKCRISISSACEDLSAALIANKLNTRMDLSNGVGFPGMULLCRGLRP
QCELOMLQKCOLESGACQEMASVLGTNPHLVELDLTGNALDELGLRLLCQGLRHP
CRRTLWMDSCGLTAKACENDYFTLGINQTLTDLYLTNNALGDTGVRLLCKRLSHP
GCKRLVLFVGMFLNMTWTHSLAALRVTPYLDIG"

misc_feature
869. .892

/notes="Region: P-loop"

misc_feature
2069. .3070

/note="Region: leucine rich repeat"

misc_feature
221. .289

/note="Region: pyrin domain"

ORIGIN

Query Match 71.5%; Score 2223; DB 9; Length 3395;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GATCCCCAGAAACCTACAGGACTATGTCGCGAGGAAATTCGCGCTCATGGAAGACCGC 348
Db |||||||
QY 590 GATCCCCAGAAACCTACAGGACTATGTCGCGAGGAAATTCGCGCTCATGGAAGACCGC 649
Db |||||||
QY 349 AATGCGCGCTAGGGAATGTGTCAACCTCAGCACCGGTACACCGGCTCTGCTGGTG 408
Db |||||||
QY 650 AATGCGCGCTAGGGAATGTGTCAACCTCAGCACCGGTACACCGGCTCTGCTGGTG 709
Db |||||||
QY 409 AAGGAGCACTCAAAACCCCATGAGGTCCAGCAGCAGCTTCTGGACACAGCGCGGACAC 468
Db |||||||
QY 710 AAGGAGCACTCAAAACCCCATGAGGTCCAGCAGCAGCTTCTGGACACAGCGCGGACAC 769
Db |||||||
QY 469 GCGAGGACCGTGGGACACCAAGGCTAGCCCATCAAGATAGACCCCTCTTTGAGCCAGAC 528
Db |||||||
QY 770 GCGAGGACCGTGGGACACCAAGGCTAGCCCATCAAGATAGACCCCTCTTTGAGCCAGAC 829
Db |||||||
QY 529 GAGGAGCGCGCGAGCACCGGACCGGTGTCATGCAAGGCGCGGACGAGATAGCAAG 588
Db |||||||
QY 830 GAGGAGCGCGCGAGCACCGGACCGGTGTCATGCAAGGCGCGGACGAGATAGCAAG 889
Db |||||||
QY 589 TCATGCTGGCACACAAAGGTGATGCTGGAGTGGGGGAGCGGGAAGCTCTTCCAAAGCAGA 648
Db |||||||
QY 890 TCATGCTGGCACACAAAGGTGATGCTGGAGTGGGGGAGCGGGAAGCTCTTCCAAAGCAGA 949
Db |||||||

649 TTTGATTATCTTCTTACATCAACTGAGGAGATGAACAGAGTGCACCGAATGCAGC 708
Db |||||
950 TTTGATTATCTTCTTACATCAACTGAGGAGATGAACAGAGTGCACCGAATGCAGC 1009
Qy |||||
709 ATGCAAGACCTCATCTTACAGTGTGCTGAGCCAGCCAGCGCTCTCCAGAGCTCATC 768
Db |||||
1010 ATGCAAGACCTCATCTTACAGTGTGCTGAGCCAGCCAGCGCTCTCCAGAGCTCATC 1069
Qy |||||
769 CGAGTCCCGAGCGCTCTTTCATCATCAACGGCTTCGATGAGCTCAAGCCCTCTTTC 828
Db |||||
1070 CGAGTCCCGAGCGCTCTTTCATCATCAACGGCTTCGATGAGCTCAAGCCCTCTTTC 1129
Qy |||||
829 CACGATCCTCAGGACCCCTGTGCTCTGCTGGAGAGAAACGGCCACCGAGCTGCTT 888
Db |||||
1130 CACGATCCTCAGGACCCCTGTGCTCTGCTGGAGAGAAACGGCCACCGAGCTGCTT 1189
Qy |||||
889 CTTAAAGCTTAATTCGAAGAAGCTGCTGCTGAGCTATCTTGTCTCATCAACACAGG 948
Db |||||
1190 CTTAAAGCTTAATTCGAAGAAGCTGCTGCTGAGCTATCTTGTCTCATCAACACAGG 1249
Qy |||||
949 CCCAGCGCTTGGAGAGCTCCACGCTGCTGAGGACCCAGCGATGTGGAGTCTG 1008
Db |||||
1250 CCCAGCGCTTGGAGAGCTCCACGCTGCTGAGGACCCAGCGATGTGGAGTCTG 1309
Qy |||||
1009 GGCTTCTCTGAGGAGAGAAAGAAATACCTTCTACAAAGTATTTCCAAATGCGAGAGCAG 1068
Db |||||
1310 GGCTTCTCTGAGGAGAGAAAGAAATACCTTCTACAAAGTATTTCCAAATGCGAGAGCAG 1369
Qy |||||
1069 GCGGCGCAAGTCTCAATTAAGTGAAGAACAGAGCTCTCTTCCCAATGTCTGCTC 1128
Db |||||
1370 GCGGCGCAAGTCTCAATTAAGTGAAGAACAGAGCTCTCTTCCCAATGTCTGCTC 1429
Qy |||||
1129 CCCCTGTGCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188
Db |||||
1430 CCCCTGTGCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1489
Qy |||||
1189 TTGAGACAGAGCTCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
Db |||||
1490 TTGAGACAGAGCTCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1549
Qy |||||
1249 CAACCCAAAGCGGGGCGCGCTCCAGGCGCTCCAGGCGCTCCAGGCGCTCCAGGCGCTCCAGG 1308
Db |||||
1550 CAACCCAAAGCGGGGCGCGCTCCAGGCGCTCCAGGCGCTCCAGGCGCTCCAGGCGCTCCAGG 1609
Qy |||||
1309 TTGGCGGAGATGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGAGCTCCGGAAG 1368
Db |||||
1610 TTGGCGGAGATGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGAGCTCCGGAAG 1669
Qy |||||
1369 CACGCGCTAGACGGGAGAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1428
Db |||||
1670 CACGCGCTAGACGGGAGAGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729
Qy |||||
1429 ATCAACTGTGAGAGTACTACAGCTTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTC 1488
Db |||||
1730 ATCAACTGTGAGAGTACTACAGCTTTCATCAAGTTCATCAAGTTCATCAAGTTCATCAAGTTC 1789
Qy |||||
1489 ATGTACTATATCTTGAAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1548
Db |||||
1790 ATGTACTATATCTTGAAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1849
Qy |||||
1549 CTGTTGACCGAGTACGCTGCTTGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1608
Db |||||
1850 CTGTTGACCGAGTACGCTGCTTGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1909
Qy |||||
1609 TTTGAGCTCTGGAACGAGAGACAGGAGCCACTGGAAGAGAGTCTCTGCTGGAAGTCTC 1668
Db |||||
1910 TTTGAGCTCTGGAACGAGAGACAGGAGCCACTGGAAGAGAGTCTCTGCTGGAAGTCTC 1969
Qy |||||
1669 TCGCGGACATCAAGATGAGCTGCTGAGTGGATCCAAAGCAAGCTCAGAGCGACGGC 1728
Db |||||
1970 TCGCGGACATCAAGATGAGCTGCTGAGTGGATCCAAAGCAAGCTCAGAGCGACGGC 2029
Qy |||||
1729 TCCACCGCTGACGAGGCTCTGAGGAGTCTTCAAGTCTGCTGAGAGATCCAGGAGGAG 1788

2030 TCCACCGCTGACGAGGCTCTCTGAGTCTTCAAGTCTTGTACGAGATCCAGGAGGAG 2089
Db |||||
1789 GAGTTTATCCAGAGGCGCTGAGCCACTTCAGGTGATCGTGGTCAGCAACATTCGCTCC 1848
Qy |||||
2090 GAGTTTATCCAGAGGCGCTGAGCCACTTCAGGTGATCGTGGTCAGCAACATTCGCTCC 2149
Db |||||
1849 AAGATGAGACATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
Qy |||||
2150 AAGATGAGACATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2209
Db |||||
1909 CACTTGTATGCGCCACTACAGCGCGGACCGGGAAGACCGCGAGGTGCTTCGACAGGA 1968
Qy |||||
2210 CACTTGTATGCGCCACTACAGCGCGGACCGGGAAGACCGCGAGGTGCTTCGACAGGA 2269
Db |||||
1969 GCGCACACGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
Qy |||||
2270 GCGCACACGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2329
Db |||||
2029 GAAATCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088
Qy |||||
2330 GAAATCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2389
Db |||||
2089 AATGCCCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
Qy |||||
2390 AATGCCCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2449
Db |||||
2149 AATGCCCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
Qy |||||
2450 AATGCCCTGGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2509
Db |||||
2209 TCTGAGCTCTCATAGCAATGAAGAAATTTGACAGGATGAGTCTCAGTGGCAACCGGCTT 2268
Qy |||||
2510 TCTGAGCTCTCATAGCAATGAAGAAATTTGACAGGATGAGTCTCAGTGGCAACCGGCTT 2569
Db |||||
2269 GGAATCCAGGATGATGCTGCTTTCAGGAGGCTGCGGCTATCCCACTGCGAGGCTGCGAG 2328
Qy |||||
2570 GGAATCCAGGATGATGCTGCTTTCAGGAGGCTGCGGCTATCCCACTGCGAGGCTGCGAG 2629
Db |||||
2329 ATGATTCAGTTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
Qy |||||
2630 ATGATTCAGTTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2689
Db |||||
2389 CTGGGACCAACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
Qy |||||
2690 CTGGGACCAACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2749
Db |||||
2449 GGCCTGAGGTTTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2508
Qy |||||
2750 GGCCTGAGGTTTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2809
Db |||||
2509 CTG 2511
Qy |||||
2810 CTG 2812
Db |||||
RESULT 9
LOCUS AY116207 3221 bp mRNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens monarch-1 splice form IV mRNA, complete cds;
ACCESSION AY116207
VERSION AY116207.1 GI:21711826
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3221)
AUTHORS Williams,K.L., Taxman,D.J., Linhoff,M.W., Reed,W. and Ting,J.P.Y.
TITLE Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat
Protein That Controls Classical and Nonclassical MHC Class I Genes
JOURNAL J. Immunol. 170 (11), 5354-5358 (2003)

PUBMED 12759408
2 (bases 1 to 3221)
REFERENCE Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
AUTHORS Direct Submission
TITLE Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
JOURNAL Road, Chapel Hill, NC 27599, USA
FEATURES Location/Qualifiers
source 1..3221
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221..2899
/note="contains alternatively spliced exons 7, 8 and 9"
/codon_start=1
/product="monarch-1 splice form IV"
/protein_id="AA075145.1"
/db_xref="GI:21711827"
/translation="MLRAGRDGLRLSTVLEBLEAVELKKFLYLGTATELGKIP
WGSMEKAGPLEMAQLLTHFGPEEAMRLALSTPERINRDLWEGQREDLVRDTPGG
PSSLGNGSTCLLESVITPRKDPOETRYDYRRKFRIMEDNRALGECVNLSHRYRL
LVKESNMQVQQLDITGRHARTVHQASPIKIEITPEDEERPEPRITVWMOA
AGIKGSLAKHVMWDAGKLFQGRFDYLFYINCRMNOSATSCMODLIFSCWPEPS
APLOELIRVPERLFIIDGDELKPSFHPDQGWCLCWEKRPTELLSLIRKLLP
ELSLITTPALEKHLRLEHPRHVELLGFSEARKEFYKIFHNAQAGQFNIVR
DNEPLFTMCFVPLVWVCTCLOQLEGGGLLROTSTTAVMYLXLLSLMQPKGAP
RUQPPNQLGSLAADGLWNLKILFEQDLRKHLDGEDVSAFLNNIPKQDINER
YYSFIHLQEFPAAMYIIDEGEGAGPDQVTRLLTEYAFSEFSLATLSRFLGL
LNEETRSLHLEKSLWCVSPHIMDLQWIOSKQSDGSLTQOGLSFFSLCYIOEE
FIQALSHVQVIVSNISKMEHWSVFLKRCRQAQVHLXYGATYSAGDELRCSA
GAHTLVQLRPERVILLDAYSEHLAALCTNPNLIELSLYRNALRGVRLCQGLRH
PNCQLNLRKRISSACEDLSAALANKNTRMDLSNGVGFPGMMLLCEGLRHP
QCLQMIQKQKQSCAGACEMSVLGTNPVHVELDLTGNALEDGLRLCQGLRHPV
CRLTLWMLFGMDLNWTHSRALALRVKPYLDIGC"
869..892
/note="Region: P-loop"
2069..2896
/note="Region: leucine rich repeat"
221..289
/note="Region: pyrim domain"
ORIGIN
Query Match 71.4%; Score 2220; DB 9; Length 3221;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 289 GATCCCCAGGAACTACAGGACTATGTCGAGGAAATTCGGGTCTATGGAAGCCGC 348
DB 590 GATCCCGAGAACTACAGGACTATGTCGAGGAAATTCGGGTCTATGGAAGCCGC 649
QY 349 AATGCGGCTAGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGCTGGT 408
DB 650 AATGCGGCTAGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGCTGGT 709
QY 409 AAGGACACTCAACCCCTCAGGTCCAGCAGCTTCTGACACAGCCCGGACAC 468
DB 710 AAGGACACTCAACCCCTCAGGTCCAGCAGCTTCTGACACAGCCCGGACAC 769
QY 469 GCGAGACCGTGGGACACCAAGGCTAGCCCCATCAAGATAGACCTCTTTGAGCAGAC 528
DB 770 GCGAGACCGTGGGACACCAAGGCTAGCCCCATCAAGATAGACCTCTTTGAGCAGAC 829
QY 529 GAGGAGCGCCGAGCCACCGCCACCGGTGATGCAAGGCGCGGAGGATAGGCAAG 588
DB 830 GAGGAGCGCCGAGCCACCGCCACCGGTGATGCAAGGCGCGGAGGATAGGCAAG 889
QY 589 TCCATGCTGACACCAAGGTGATGCTGAGTGGCGGAGGAGCTCTTCCAGGACGA 648
DB 890 TCCATGCTGACACCAAGGTGATGCTGAGTGGCGGAGGAGCTCTTCCAGGACGA 949
QY 649 TTTGATTATCTTCTACATCAATCAGGAGATGAACAGAGTGCACGAGGATGCAGC 708

DB 950 TTTGATTATCTTCTACATCAATGCGAGGAGATGAACAGAGTGCACGGAATGCAGC 1009
QY 709 ATGCAAGACCTCATCTTTAGCTGTGCGCTGAGCCAGCGCGCTCTTCCAGGAGCTCATC 768
DB 1010 ATGCAAGACCTCATCTTTAGCTGTGCGCTGAGCCAGCGCGCTCTTCCAGGAGCTCATC 1069
QY 769 CGAGTTCCGAGCGCTCTTTTCAATCATCGAGCGCTTTCAGATGAGCTCAAGCCCTTTTC 828
DB 1070 CGAGTTCCGAGCGCTCTTTTCAATCATCGAGCGCTTTCAGATGAGCTCAAGCCCTTTTC 1129
QY 829 CAGATCTCAGGAGCGCTGTGCTGTGCGGAGGAGAAACGCGCCAGCGAGCTGCTT 888
DB 1130 CAGATCTCAGGAGCGCTGTGCTGTGCGGAGGAGAAACGCGCCAGCGAGCTGCTT 1189
QY 889 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 948
DB 1190 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 1249
QY 949 CCACGCGTTTGAGAGAGCTCCACCGTCTCTGAGACACCCAGCGCATGTGGAGATCCTG 1008
DB 1250 CCACGCGTTTGAGAGAGCTCCACCGTCTCTGAGACACCCAGCGCATGTGGAGATCCTG 1309
QY 1009 GGCTTCTCAGGACGAGAAAGGAATATCTTCTACAAGTATTTCCACAATGACAGACG 1068
DB 1310 GGCTTCTCAGGACGAGAAAGGAATATCTTCTACAAGTATTTCCACAATGACAGACG 1369
QY 1069 GCGGCGCAAGCTTCAATTAAGTGAAGGACCAACAGAGCTCTTTCACCATGTGCTCGTC 1128
DB 1370 GCGGCGCAAGCTTCAATTAAGTGAAGGACCAACAGAGCTCTTTCACCATGTGCTCGTC 1429
QY 1129 CCCTGTGTGCTGGGTGTGTGTAAGTCTGAGCAGAGCTGAGAGGTGGGGGCTG 1188
DB 1430 CCCTGTGTGCTGGGTGTGTGTAAGTCTGAGCAGAGCTGAGAGGTGGGGGCTG 1489
QY 1189 TTGAGACAGAGCTCCAGGACCAACCTGAGTGTGATGCTTACCTGTGAGTCTCATG 1248
DB 1490 TTGAGACAGAGCTCCAGGACCAACCTGAGTGTGATGCTTACCTGTGAGTCTCATG 1549
QY 1249 CAACCCAAAGCGGGGCGCGCGCTCCAGCGCCCGCCAGCCAGAGAGGAGGTGTGCTCC 1308
DB 1550 CAACCCAAAGCGGGGCGCGCGCTCCAGCGCCCGCCAGCCAGAGAGGAGGTGTGCTCC 1609
QY 1309 TTGCGCGCAGATGGGTCTGGAAATCAGAAATCTTATTTGAGGAGAGGAGCTTCCGGAAG 1368
DB 1610 TTGCGCGCAGATGGGTCTGGAAATCAGAAATCTTATTTGAGGAGAGGAGCTTCCGGAAG 1669
QY 1369 CAGCGCTTAGACGCGGAGACGCTCTGCTCTCTCAACATGAACATCTTCCAGAGAGAC 1428
DB 1670 CAGCGCTTAGACGCGGAGACGCTCTGCTCTCTCAACATGAACATCTTCCAGAGAGAC 1729
QY 1429 ATCAACTGTGAGAGTACTACAGCTTCACTCACTTTCAGGATTTCCAGGAATCTTTGACGT 1488
DB 1730 ATCAACTGTGAGAGTACTACAGCTTCACTCACTTTCAGGATTTCCAGGAATCTTTGACGT 1789
QY 1489 ATGTAATATCTTGAACGAGGAGGCGGGGAGGCGGCGGAGGCGGAGGAGGAGGAGGAG 1548
DB 1790 ATGTAATATCTTGAACGAGGAGGCGGGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAG 1849
QY 1549 CTGTTACCGAGTACGCGTCTTCTGAAAGAGGAGTCTCTGCACTCAGGAGGAGGAGGAGGAG 1608
DB 1850 CTGTTACCGAGTACGCGTCTTCTGAAAGAGGAGTCTCTGCACTCAGGAGGAGGAGGAGGAG 1909
QY 1609 TTTGGAATCTCTGAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
DB 1910 TTTGGAATCTCTGAAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1969
QY 1669 TCGCCGACATCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1728
DB 1970 TCGCCGACATCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2029
QY 1729 TCCACCGCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1788
DB 2030 TCCACCGCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2089

1474 CAACCCAGCGGGGCGGCTCCAGCCGCCACCCACAGAGAGGTTGTGCTCC 1533
1309 TTGGGGGAGATGGCTCTGGAATCAGAAAATCTATTGAGGAGCAGGACTCCGGAAG 1368
1534 TTGGGGGAGATGGCTCTGGAATCAGAAAATCTATTGAGGAGCAGGACTCCGGAAG 1593
1369 CACGGCTTAGACGGGAGAGAGCTCTGCGCTTCTCAACATGAACATCTTCAGAGGAC 1428
1594 CACGGCTTAGACGGGAGAGAGCTCTGCGCTTCTCAACATGAACATCTTCAGAGGAC 1653
1429 ATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTGAGCT 1488
1654 ATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTGAGCT 1713
1489 ATGTAATATATCTTGGAGAGGGGAGGCGGGGAGGCGCCAGACCGAGAGCTGACCAAG 1548
1714 ATGTAATATATCTTGGAGAGGGGAGGCGGGGAGGCGCCAGACCGAGAGCTGACCAAG 1773
1549 CTGTTGACCGAGTACGGGCTTTCTGAAGGAGCTTCTGGCACTCACCAGCGCTTCTCTG 1608
1774 CTGTTGACCGAGTACGGGCTTTCTGAAGGAGCTTCTGGCACTCACCAGCGCTTCTCTG 1833
1609 TTTGGAATCTCTGAACGAGGAGCAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC 1668
1834 TTTGGAATCTCTGAACGAGGAGCAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC 1893
1669 TCGCGGACATCAGATGGAGCTTGTGAGTGGATCCAAAGCAAGCTCAGAGCGGC 1728
1894 TCGCGGACATCAGATGGAGCTTGTGAGTGGATCCAAAGCAAGCTCAGAGCGGC 1953
1729 TCCACCTGTCAGGAGGCTCTCTGAGTCTTCTGAGTCTTGTGAGAGTCCAGAGGAG 1788
1954 TCCACCTGTCAGGAGGCTCTCTGAGTCTTCTGAGTCTTGTGAGAGTCCAGAGGAG 2013
1789 GAGTTTATCCAGCAGGCGCTTCTGAGCCTTCCAGGTGATCTGCTGTCAGCAATGCTCTCC 1848
2014 GAGTTTATCCAGCAGGCGCTTCTGAGCCTTCCAGGTGATCTGCTGTCAGCAATGCTCTCC 2073
1849 AAGATGAGCAGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
2074 AAGATGAGCAGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2133
1909 CACTTGTATGCGGCCACTACAGCGGAGCGGAGAGCGGAGCGGAGGCTCTCGCAGGA 1968
2134 CACTTGTATGCGGCCACTACAGCGGAGCGGAGAGCGGAGGCTCTCGCAGGA 2193
1969 GCGCACACGCTGTGTGTCAGCTCAGACCGAGAGGAGCGGCTTCTGCTGAGCGCTACAGT 2028
2194 GCGCACACGCTGTGTGTCAGCTCAGACCGAGAGGAGCGGCTTCTGCTGAGCGCTACAGT 2250
2029 GAACATCTGCGAGCGGCGCTTGTGCAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 2088
2251 GAACATCTGCGAGCGGCGCTTGTGCAATCCAAATCCAAATCCAAATCCAAATCCAAATCC 2310
2089 AATGCTCTGGGAGCGGCGGCTGAGCTCTCTGTCAGAGTCTCAGACCGGAGGCTGCTG 2148
2311 AATGCTCTGGGAGCGGCGGCTGAGCTCTCTGTCAGAGTCTCAGACCGGAGGCTGCTG 2370
2149 AACTTTAGAACTGAGGTGAGAGGTGCGGATCTTCCAGCTCAGCTCAGGAGGAGCTC 2208
2371 AACTTTAGAACTGAGGTGAGAGGTGCGGATCTTCCAGCTCAGCTCAGGAGGAGCTC 2430
2209 TCTGAGCTCTCATAGCCATAGAAATTTGACAGAGATGATCTCAGTGGCAACGGGCTT 2268
2431 TCTGAGCTCTCATAGCCATAGAAATTTGACAGAGATGATCTCAGTGGCAACGGGCTT 2490
2269 GGATTTCCAGGATGATGCTGCTTTTCCAGGAGGCTGCGGATCCCGAGTCCAGGCTGCG 2328
2491 GGATTTCCAGGATGATGCTGCTTTTCCAGGAGGCTGCGGATCCCGAGTCCAGGCTGCG 2550
2329 ATGATTCAGTTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2388
2551 ATGATTCAGTTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2610

QY 2389 CTCGSCAACCAACCCACATCTCTGTTGAGTTGACCTGACAGAAATGCACTGGAGATTG 2448
Db 2611 CTTGGSCAACCAACCCACATCTCTGTTGAGTTGACCTGACAGAAATGCACTGGAGATTG 2670
QY 2449 GGCCTGAGGTTTACTATGTCAGGAGCTGAGGACCCAGTCTCAGACTACGAGCTTTGTGG 2508
Db 2671 GGCCTGAGGTTTACTATGTCAGGAGCTGAGGACCCAGTCTCAGACTACGAGCTTTGTGG 2730
QY 2509 CTGAAGATCTGCCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
Db 2731 CTGAAGATCTGCCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2790
QY 2569 AACCAGAGCTTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2628
Db 2791 AACCAGAGCTTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2850
QY 2629 CTGCTGTGTGAGGCGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGG 2679
Db 2851 CTGCTGTGTGAGGCGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGG 2901

RESULT 11
AK095460
LOCUS 3466 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ38141 fis, clone D90ST2002673, weakly similar to Homo sapiens caspase recruitment domain protein 7 mRNA.
ACCESSION AK095460
VERSION AK095460.1 GI:21754717
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Ohshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3466)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1. 3466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="D90ST2002673"
/cell_type="CD34+ Cells"
/clone_lib="D90ST2"
/note="cloning vector: pME18SFL3-mRNA from CD34+ cells after 9-days ODF induction.-primary culture, CD34+ Cells"

ORIGIN
Query Match 64.2%; Score 1995; DB 9; Length 3466;

Best Local Similarity 99.7%; Pred. No. 0;										
Matches 2385; Conservative 0; Mismatches 3; Indels 3; Gaps 1;										
QY	289	GATCCCCAGGAAACCTACAGGGA	TA	TG	TC	CG	AG	AA	TTCCGGCTCATGGAAGACGC	348
DB	514	GATCCCCAGGAAACCTACAGGGA	TA	TG	TC	CG	AG	AA	TTCCGGCTCATGGAAGACGC	573
QY	349	AATCGCGCTTAGGGAATGTGTCA	CT	CA	CT	CA	CG	CA	CGGTACACCGGCTCCTGCTGTG	408
DB	574	AATCGCGCTTAGGGAATGTGTCA	CT	CA	CT	CA	CG	CA	CGGTACACCGGCTCCTGCTGTG	633
QY	409	AAGGAGCACTCAAAACCCATGAG	GT	TC	AG	CA	GT	TTCTGGACACAGGCGGGGACAC	468	
DB	634	AAGGAGCACTCAAAACCCATGAG	GT	TC	AG	CA	GT	TTCTGGACACAGGCGGGGACAC	693	
QY	469	GCAGGACCGTGGGACACCAAGG	CT	AG	CC	CA	AT	CAAGATAGAGACCCCTCTTTGAGCCAGAC	528	
DB	694	GCAGGACCGTGGGACACCAAGG	CT	AG	CC	CA	AT	CAAGATAGAGACCCCTCTTTGAGCCAGAC	753	
QY	529	GAGGAGCGCCGAGCCAGCGCA	CG	CG	CA	CG	CG	CGGTATGCAAGCGCGCAGGATAGGCAAG	588	
DB	754	GAGGAGCGCCGAGCCAGCGCA	CG	CG	CA	CG	CG	CGGTATGCAAGCGCGCAGGATAGGCAAG	813	
QY	589	TCCATGCTGGCACACAAGGTGAT	GT	GA	CT	GG	CG	GACCGGAAGCTCTTCCAAAGGCAGA	648	
DB	814	TCCATGCTGGCACACAAGGTGAT	GT	GA	CT	GG	CG	GACCGGAAGCTCTTCCAAAGGCAGA	873	
QY	649	TTTGATTAATCTCTTACATCAAT	CT	CA	T	CA	T	CA	TAACTGCGGGAGATGAACCGAGATGCGACGAATGCGAC	708
DB	874	TTTGATTAATCTCTTACATCAAT	CT	CA	T	CA	T	CA	TAACTGCGGGAGATGAACCGAGATGCGACGAATGCGAC	933
QY	709	ATGCAAGACCTCATCTTCAAGCT	GT	CG	CT	GG	CT	GG	CGCCAGCGGCTCTCCAGGAGCTCATC	768
DB	934	ATGCAAGACCTCATCTTCAAGCT	GT	CG	CT	GG	CT	GG	CGCCAGCGGCTCTCCAGGAGCTCATC	993
QY	769	CGAGTTCCGAGCGGCTCTTTTCA	T	CA	T	CA	T	CA	ATCGACGGCTTCGATGAGCTCAAGCTCTTCTTC	828
DB	994	CGAGTTCCGAGCGGCTCTTTTCA	T	CA	T	CA	T	CA	ATCGACGGCTTCGATGAGCTCAAGCTCTTCTTC	1053
QY	829	CACGATCTTCAGGACCCCTGTGC	CT	GT	CG	AG	GAGAAACCGGCCACCGAGCTGCTT	888		
DB	1054	CACGATCTTCAGGACCCCTGTGC	CT	GT	CG	AG	GAGAAACCGGCCACCGAGCTGCTT	1113		
QY	889	CTTAAACAGCTTAATTCGGAAGA	AGCTGCTCCCTGAGCTATCT	TTTGTCTCATCAACCAACGC	948					
DB	1114	CTTAAACAGCTTAATTCGGAAGA	AGCTGCTCCCTGAGCTATCT	TTTGTCTCATCAACCAACGC	1173					
QY	949	CCACGGCTTTGAGAGAGCTCCAC	CGTCTGCTGGAGCACCCACGAGCATGTGGAGATCCCTG	1008						
DB	1174	CCACGGCTTTGAGAGAGCTCCAC	CGTCTGCTGGAGCACCCACGAGCATGTGGAGATCCCTG	1233						
QY	1009	GGCTTCTCTGAGCGAGAAAGGAG	GAATACTTCTACAAGTATTTCCAAATGCGAGAGCG	1068						
DB	1234	GGCTTCTCTGAGCGAGAAAGGAG	GAATACTTCTACAAGTATTTCCAAATGCGAGAGCG	1293						
QY	1069	GGGGCCAAAGTCTTCAATTACGT	GAGGACAACGAGCTCTCTTCAACATGTGCTTCGTC	1128						
DB	1294	GGGGCCAAAGTCTTCAATTACGT	GAGGACAACGAGCTCTCTTCAACATGTGCTTCGTC	1353						
QY	1129	CCCCTGGTGTGCTGGGTGGTGTG	TACCTGCTCCAGCAGCAGCTGGAGGCTGGGGGGCTG	1188						
DB	1354	CCCCTGGTGTGCTGGGTGGTGTG	TACCTGCTCCAGCAGCAGCTGGAGGCTGGGGGGCTG	1413						
QY	1189	TTGAGACAGACGTCCAGGACCA	CCACTGACGTGTACATGCTCTACCTGCTGAGTCTGATG	1248						
DB	1414	TTGAGACAGACGTCCAGGACCA	CCACTGACGTGTACATGCTCTACCTGCTGAGTCTGATG	1473						
QY	1249	CAACCCAGCCGGGGCCCCGGCTC	ACGCCCCCACCAACAGAGAGGGTTGTGCTCC	1308						
DB	1474	CAACCCAGCCGGGGCCCCGGCTC	ACGCCCCCACCAACAGAGAGGGTTGTGCTCC	1533						
QY	1309	TTGGCGGACAGATGGGCTCTG	GAATCAGAAAATCTTATTTGAGGAGCAGGACCTCCGGAAG	1368						

Db	1534	TTGGCGGCGAGATGGGCTCTGGGAATCAGAAATCTTATTTTGAGGAGCAGGACCTCCGGAAG	1593
Qy	1369	CACGGCTTAGACGGGGAAGACGCTCTCTGCTTCTCTCAACATGAACATCTTCCAGAAAGGAC	1428
Db	1594	CACGGCTTAGACGGGGAAGACGCTCTCTGCTTCTCTCAACATGAACATCTTCCAGAAAGGAC	1653
Qy	1429	ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTTTCAGCT	1488
Db	1654	ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTTTCAGCT	1713
Qy	1489	ATGTACTATATCTCTGACGAGGGGAGGGCGGGCAGGCCAGCCAGACCAAGCAAGCTGACCAAG	1548
Db	1714	ATGTACTATATCTCTGACGAGGGGAGGGCGGGCAGGCCAGACCAAGCAAGCTGACCAAG	1773
Qy	1549	CTGTTGACCGAGTACCGCTTTTCTGAAAGAGAGCTTCTTGCCACTCACACGCGCTTCTCTG	1608
Db	1774	CTGTTGACCGAGTACCGCTTTTCTGAAAGAGAGCTTCTTGCCACTCACACGCGCTTCTCTG	1833
Qy	1609	TTTGGACTCTGAAACAGAGAGACGAGGACCCAGGAGCCACTGGAGAGAGTCTCTGCTGGAGGTC	1668
Db	1834	TTTGGACTCTGAAACAGAGAGACGAGGAGCCACTGGAGAGAGTCTCTGCTGGAGGTC	1893
Qy	1669	TCGCCGACATCAAGATGACCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGACGGC	1728
Db	1894	TCGCCGACATCAAGATGACCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGACGGC	1953
Qy	1729	TCCACCTGACAGAGGCGCTCCTTTGGAGTTCTTTCAGCTGCTTTGACGAGTCCAGAGGAG	1788
Db	1954	TCCACCTGACAGAGGCGCTCCTTTGGAGTTCTTTCAGCTGCTTTGACGAGTCCAGAGGAG	2013
Qy	1789	GAGTTTATCAGCAGGCGCTTGAGCAGCTTCCAGGTGATCGTGTGTCAGCAACATTTGCTCC	1848
Db	2014	GAGTTTATCAGCAGGCGCTTGAGCAGCTTCCAGGTGATCGTGTGTCAGCAACATTTGCTCC	2073
Qy	1849	AAGATGAGCACATGCTCTCTGTTCTGTAAGCGCTCAGAGAGCGCCAGGCTGCTG	1908
Db	2074	AAGATGAGCACATGCTCTCTGTTCTGTAAGCGCTCAGAGAGCGCCAGGCTGCTG	2133
Qy	1909	CACTTGTATGGCGCCACTACAGCGCGGACGGGGAAGACCGCGAGAGTGTCTCCGACGA	1968
Db	2134	CACTTGTATGGCGCCACTACAGCGCGGACGGGGAAGACCGCGAGAGTGTCTCCGACGA	2193
Qy	1969	GGCGACACGCTTTGGTGAGCTCAGACCGAGAGAGACCGTTCTGTCTGGAGCGCTACAGT	2028
Db	2194	GGCGACACGCTTTGGTGAGCTCAGACCGAGAGAGACCGTTCTGTCTGGAGCGCTACAGT	2250
Qy	2029	GAAATCTGGCAGCGCGCTGTGCACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA	2088
Db	2251	GAAATCTGGCAGCGCGCTGTGCACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA	2310
Qy	2089	AATGCCCTGGCAGCGCGGGGTGAAGCTGTCTGTCAAGACTCAGACACCCCACTGC	2148
Db	2311	AATGCCCTGGCAGCGCGGGGTGAAGCTGTCTGTCAAGACTCAGACACCCCACTGC	2370
Qy	2149	AAACTTCAGAACTGAGGCTGAAGGTGGCGCATCTCAGCTCAGCCTGGGAGGACCTC	2208
Db	2371	AAACTTCAGAACTGAGGCTGAAGAGGTGGCGCATCTCCAAGCTCAGCCTGGGAGGACCTC	2430
Qy	2209	TCTGAGCTCTCATAGCCATAGAAATTTGACAGAGATGGATCTCAGTGCGCAAGCGGTT	2268
Db	2431	TCTGAGCTCTCATAGCCATAGAAATTTGACAGAGATGGATCTCAGTGCGCAAGCGGTT	2490
Qy	2269	GGATTTCCAGGACATGATGCTCTTTGCGAGGGCGCTCGCGCATCCCAAGTGCAGGCTCGAG	2328
Db	2491	GGATTTCCAGGACATGATGCTCTTTGCGAGGGCGCTCGCGCATCCCAATGCAGGCTCGAG	2550
Qy	2329	ATGATTTCAAGTTGAGGAAGTGTCACTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTG	2388
Db	2551	ATGATTTCAAGTTGAGGAAGTGTCACTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTG	2610
Qy	2389	CTCGGACCAACCAATCTCGTTGATTTGACCTGACAGGAATGCACTGGAGATTTG	2448
Db	2611	CTTGSCACCAACCAATCTCGTTGATTTGATTTGACCTGACAGGAATGCACTGGAGATTTG	2670

Db	838	GATCTCAGTCGGCAACGGCGTTGGANTCCCAAGGCATGANGTCTTTTCGAGGGCGCTCGCGG	897
Qy	2308	CATCCCAAGTCAGGCTGCAGATGATTCAGTTGAGGGAAGTGTCTCAGCTGGAGTCCGGGCT	2367
Db	898	CATCCCAATGCGAGGCTGCAGATGATTCAGTTGAGGGAAGTGTCTCAGCTGGAGTCCGGGCT	957
Qy	2368	TGTCAGAGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGGTTGAGTTGGACCTGACA	2427
Db	958	TGTCAGAGAGATGGCTTCTGTGCTTGGCACCAACCCACATCTGGTTGAGTTGGACCTGACA	1017
Qy	2428	GGAAATGCACTGGAGGATTTTGGGCTCAGGTTACTATGCCAGGAGCTGAGGCCACCCAGTC	2487
Db	1018	GGAAATGCACTGGAGGATTTTGGGCTCAGGTTACTATGCCAGGAGCTGAGGCCACCCAGTC	1077
Qy	2488	TGCAGACTACGACTTTTGTGGCTGAAGATCTGCCGCTCACTCTGCTGCTCTGTGACGAG	2547
Db	1078	TGCAGACTACGACTTTTGTGGCTGAAGATCTGCCGCTCACTCTGCTGCTCTGTGACGAG	1137
Qy	2548	CTGGCTCAACTCTCAGTGTGAACCAAGAGCTCTGAGAGAGCTGAGAGCTGAGAGCTGAGATGAG	2607
Db	1138	CTGGCTCAACTCTCAGTGTGAACCAAGAGCTCTGAGAGAGCTGAGAGCTGAGAGCTGAGATGAG	1197
Qy	2608	CTGGGGAACCTCGGGGTGCTGCTGCTGTGTGAGGGCCTCAGGCATCCAGCTGCAAGCTC	2667
Db	1198	CTGGGGAACCTCGGGGTGCTGCTGCTGTGTGAGGGCCTCAGGCATCCAGCTGCAAGCTC	1257
Qy	2668	CAGACCTCGGTTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTG	2727
Db	1258	CAGACCTCGGTTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTG	1317
Qy	2728	GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTGATGTTTCAACGACCTGGGAGAC	2787
Db	1318	GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTGATGTTTCAACGACCTGGGAGAC	1377
Qy	2788	TGGGGCTGTGGTTGTGGCTGAGGGGCTGCAACATCCCGCTTCGAGACTCCAGAACTG	2847
Db	1378	TGGGGCTGTGGTTGTGGCTGAGGGGCTGCAACATCCCGCTTCGAGACTCCAGAACTG	1437
Qy	2848	TGGCTGATAGCTGTGSCCTCAGAGCCAGGCTGTGAGNACTTTTACTTTCACCCCTGGGG	2907
Db	1438	TGGCTGATAGCTGTGSCCTCAGAGCCAGGCTGTGAGNACTTTTACTTTCACCCCTGGGG	1497
Qy	2908	ATCAACAGACCTTGACCGACCTTTACTTGACCAACACGCCCTTAGGGGACACAGGTGTC	2967
Db	1498	ATCAACAGACCTTGACCGACCTTTACTTGACCAACACGCCCTTAGGGGACACAGGTGTC	1557
Qy	2968	CGACTCTTTGCAAGGGGCTGAGGCATCTGGGCTGAAACTCCGAGTCTCTGGTTATTT	3027
Db	1558	CGACTCTTTGCAAGGGGCTGAGGCATCTGGGCTGAAACTCCGAGTCTCTGGTTATTT	1617
Qy	3028	GGGATGGACTGTAATAAATGACCCAAGTGTGGCAGCGCTTCGAGTAAACAAACCT	3087
Db	1618	GGGATGGACTGTAATAAATGACCCAAGTGTGGCAGCGCTTCGAGTAAACAAACCT	1677
Qy	3088	TATTTGGACATTTGGCTGCTGA	3108
Db	1678	TATTTGGACATTTGGCTGCTGA	1698

RESULT 14

AX684291	AX684291	4931 bp	DNA	linear	PAT 29-MAR-2003
LOCUS	Sequence 1 from Patent WO02052011.				
DEFINITION	AX684291				
ACCESSION	AX684291				
VERSION	AX684291.1	GI:29371161			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Feder, J., Ramanathan, C. and Mintier, G.				
TITLE	Human leucine-rich repeat containing protein, hlrbm1, expressed				

Db 1974 AATCTCTATTGAGGAGCAGGACCTCCGAGACACGCGCTAGACGGGAGAGGCTCTCT 2033
QY 1396 GCCTTCTCTAACATGAACATCTCTCCAGAGGACATCAACTGTGAGAGG 1443
Db 2034 GCCTTCTCTAACATGAACATCTCTCCAGAGGACATCAACTGTGAGAGG 2081

RESULT 15
AF231021 2494 bp mRNA linear PRI 02-MAR-2001
LOCUS Homo sapiens leucine-rich-repeat protein RNO2 mRNA, complete cds.
DEFINITION AF231021
ACCESSION AF231021
VERSION AF231021.1 GI:13182796
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2494)
Shami,P.J., Kanai,N., Wang,L.Y., Vreeke,T.M. and Parker,C.J.
IDENTIFICATION AND CHARACTERIZATION OF A NOVEL GENE THAT IS
UPREGULATED IN LEUKEMIA CELLS BY NITRIC OXIDE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2494)
Shami,P.J., Kanai,N., Wang,L.Y., Vreeke,T.M. and Parker,C.J.
AUTHORS Direct Submission
TITLE Submitted (02-FEB-2000) Medicine, University of Utah and SLC VA
JOURNAL Medical Centers, Box 151M, 500 Foothill Boulevard, Salt Lake City,
UT 84148, USA

FEATURES
source Location/Qualifiers
1..2494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1309..2172
/note="similar to the Homo sapiens ribonuclease/angiogenin
inhibitor; expressed in hematopoietic cells; upregulated
by nitric oxide in leukemia cells"
/codon_start=1
/product="leucine-rich-repeat protein RNO2"
/protein_id="AAK14942.1"
/db_xref="GI:13182797"
/translation="MSQAWHTSVSPATQEAAGAGGATGATCTCAGTGGCAACGGCGTTGGATTCCAGGCGATGA 2284
LSAALIANKNLTRMDLNSGVGPGFMILLCEGLRHPQCRLOMIOIRKCOLESAGAOEM
ASVLGTHPLVELDITGNALDELGRLLCCGRHPVCKRTLWIKI CRLFRAACDELA
STLSVNSQLRELDLSINELDGLVLLCEGLRHPTCKLQTLRLDSCGLTAKACENLYF
TLGINQTLDLYLTNNALGDTGVRLLCRLSHPGCKLRVLWLPFGMDLNKMTSHRLAL
RVTKPYLDIGC"

ORIGIN
Query Match 16.6%; Score 515; DB 9; Length 2494;
Best Local Similarity 100.0%; Pred. No. 1.4e-277;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2165 GCCTGAAGAGGTGCGCGATCTCCAGCTCAGCTGCGAGACCTCTCTGAGCTCTCATAG 2224
Db 1400 GCCTGAAGAGGTGCGCGATCTCCAGCTCAGCTGCGAGACCTCTCTGAGCTCTCATAG 1459
QY 2225 CCAATAGAATTTGACAAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCAGGCGATGA 2284
Db 1460 CCAATAGAATTTGACAAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCAGGCGATGA 1519
QY 2285 TGCTGCTTTGCGAGGCGCTCGCGGATCCCGAGTCGAGGCTGAGATGATTGAGTTGAGGA 2344
Db 1520 TGCTGCTTTGCGAGGCGCTCGCGGATCCCGAGTCGAGGCTGAGATGATTGAGTTGAGGA 1579
QY 2345 AGTGTGAGCTGAGTCCGGGGCTTCTCAGGAGATGGCTTCTGCTCGGCACCAACCCAC 2404
Db 1580 AGTGTGAGCTGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGCTCGGCACCAACCCAC 1639
QY 2405 ATCTGGTTGAGTTGGACCTGCAGGAAATGCACTGAGGATTGGCGCTGAGGTTACTAT 2464

Db 1640 ATCTGTTGAGTTGGACCTGCACAGGAAATGCACCTGGAGGATTTGGGCTGAGGTTACTAT 1699
QY 2465 GCCAGGAGCTGAGGCGACCCAGTCTGCGAGACTAGCGACTTTTGTGGCTGAAGATCTGCCGCC 2524
Db 1700 GCCAGGAGCTGAGGCGACCCAGTCTGCGAGACTAGCGACTTTTGTGGCTGAAGATCTGCCGCC 1759
QY 2525 TCACTGCTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAG 2584
Db 1760 TCACTGCTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAG 1819
QY 2585 AGCTGGACCTGAGGCTGAATGAGCTGGGGGACCTCGGGGTGCTGCTGCTGTGTGAGGGCC 2644
Db 1820 AGCTGGACCTGAGGCTGAATGAGCTGGGGGACCTCGGGGTGCTGCTGCTGTGTGAGGGCC 1879
QY 2645 TCAGGCATCCACAGCTGCAAGCTCCAGACCTCGCG 2679
Db 1880 TCAGGCATCCACAGCTGCAAGCTCCAGACCTCGCG 1914

Search completed: August 8, 2004, 14:03:09
Job time : 11833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 07:04:50 ; Search time 1132 Seconds
(without alignments)
11663.781 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctacgaacgcaggcag.....atttggacattggctgtgga 3108

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3108	100.0	3108	6 ADE36416	Human PAA
2	3057	98.4	3300	6 AAL47129	Pyrim dom
3	2875	79.6	3186	6 AAL44363	Human PFR
4	2096	67.4	3306	9 ADC30316	Human nov
5	1356	43.6	2158	6 ABS78719	Human cDN
6	1212	39.0	1683	7 ACD03624	Novel hum
7	1116	35.9	1800	7 ACD03623	Novel hum
8	768	24.7	4931	6 ABL59333	Novel hum
9	520	16.7	591	6 ADC32201	Human nov
10	385	12.4	487	8 ACH36225	Human end
11	346	11.1	479	3 AAC76566	Human ORF
12	289	9.3	594	4 AAK37554	Human bon
13	289	9.3	3306	9 ADC30316	Human nov
14	282	9.1	282	4 AAK50424	Human bon
15	254	8.2	506	5 AAS68757	Human bon
16	173	5.6	515	4 AAK37572	Human bon
17	171	5.5	171	4 AAK50441	Human bon
18	152	4.9	289	4 AAK67173	Human imm
19	148	4.8	299	6 AAL44361	Human PFR
20	143	4.6	308	6 ABL59334	Expressed
21	136	4.4	352	4 AAK56010	Human imm
22	69	2.2	80	6 ABL59337	PCR prime
23	52	1.7	919	5 AAS68285	DNA encod

c	24	34	1.1	2149	5 AAS79967	Aas79967 DNA encod
	25	26	0.8	26	7 ACD03668	Novel hum
	26	26	0.8	39	6 ABL59339	PCR prime
c	27	25	0.8	37	6 ABL59340	PCR prime
c	28	25	0.8	37	6 ABL59342	PCR prime
	29	24	0.8	24	6 AAD33528	Novel hum
	30	24	0.8	38	6 ABL59341	Novel hum
	31	22	0.7	22	7 ACD03667	Novel hum
c	32	22	0.7	22	7 ACD03669	Novel hum
	33	21	0.7	21	6 ABL59335	PCR prime
c	34	21	0.7	21	6 ABL59336	PCR prime
c	35	21	0.7	25	7 ACD03674	Novel hum
c	36	21	0.7	25	7 ACD03677	Novel hum
c	37	21	0.7	25	7 ACD03680	Novel hum
	38	21	0.7	455	8 ACH40376	Human fce
	39	21	0.7	467	4 AAL11592	Probe #15
	40	21	0.7	467	4 ABA53281	Human fce
	41	21	0.7	467	4 AAL32884	Probe #15
	42	21	0.7	467	4 ABA42858	Human bre
	43	21	0.7	467	4 ABA23055	Probe #15
	44	21	0.7	467	4 AAK26984	Human bon
	45	21	0.7	467	4 AAK01539	Human bra

ALIGNMENTS

RESULT 1

AD36416
ID ADE36416 standard; cDNA; 3108 BP.

AC ADE36416;

DT 29-JAN-2004 (first entry)

DE Human PAA and nucleotide binding protein PAN6 cDNA.

KW cytostatic; immunosuppressive; vulnary; antiinflammatory; vasotropic;
KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiac;
KW antiparkinsonian; nortropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAA domain containing polypeptide;
KW PAA and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAA domain family; human; PAN6; gene; ss.

OS Homo sapiens.

XX US2003077699-A1.

XX 24-APR-2003.

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

XX 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (CHUZ/) CHU Z.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (ARIZ/) ARIZA M E.

PT involving impaired signal transduction, particularly inflammation, also
PT proteins and antibodies.

PS Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where
CC the proteins contain at least one PYD (pyrin) domain. These can be used
CC to treat diseases associated with impaired intracellular signal
CC transduction, particularly inflammation such as psoriasis,
CC arteriosclerosis, bacterial or viral infections (particularly meningitis,
CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
CC and Parkinson's diseases. The present sequence is a coding sequence of
CC the invention

XX Sequence 3300 BP; 726 A; 943 C; 979 G; 652 T; 0 U; 0 Other;

Query Match 98.4%; Score 3057; DB 6; Length 3300;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTACGAAACCGCAGGAGGAGCGCTCTGTCGCTGTCCACCTACTTGGAAAGACTC	60
DB	1	ATGCTACGAAACCGCAGGAGGAGCGCTCTGTCGCTGTCCACCTACTTGGAAAGACTC	60
QY	61	GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGAGAA	120
DB	61	GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGAGAA	120
QY	121	GGCAAGATCCCTGGGGAAGCATGAGAGAGCCGGTCCCTGGAAATGGCCCGACGTGCTC	180
DB	121	GGCAAGATCCCTGGGGAAGCATGAGAGAGCCGGTCCCTGGAAATGGCCCGACGTGCTC	180
QY	181	ATCACCCACATTCGGGCCAGAGAGGCGCTGAGGTGGCTCTCAGCACCTTTGAGCGGATA	240
DB	181	ATCACCCACATTCGGGCCAGAGAGGCGCTGAGGTGGCTCTCAGCACCTTTGAGCGGATA	240
QY	241	AACAGGAGGACCTGTGGGAGAGAGACAGAGAGGACCTGGTGAGGGATCCCCAGGAA	300
DB	241	AACAGGAGGACCTGTGGGAGAGAGACAGAGAGGACCTGGTGAGGGATCCCCAGGAA	300
QY	301	ACCTACAGGAGCTATGTCGAGGAAATTCGGGCTCATGGAAGACCGCAATGCGCGCTA	360
DB	301	ACCTACAGGAGCTATGTCGAGGAAATTCGGGCTCATGGAAGACCGCAATGCGCGCTA	360
QY	361	GGGGAATGTGCAACTCAGCCACCGGTACACCGGCTCCTGCTGTCGAAGAGCACTCA	420
DB	361	GGGGAATGTGCAACTCAGCCACCGGTACACCGGCTCCTGCTGTCGAAGAGCACTCA	420
QY	421	AACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCGCGGGACACGCGAGGACCGTG	480
DB	421	AACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCGCGGGACACGCGAGGACCGTG	480
QY	481	GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCACAGAGAGGCGCCC	540
DB	481	GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCACAGAGAGGCGCCC	540
QY	541	GAGCCACCGCAGCGTGTATGCAAGCGCGGACAGGATAGGCAAGTCCATGCTGGCA	600
DB	541	GAGCCACCGCAGCGTGTATGCAAGCGCGGACAGGATAGGCAAGTCCATGCTGGCA	600
QY	601	CACAGGTGATGCTGGAGCTGGCGGACGGGAAGCTCTTCCAAAGGAGATTTGATATCTC	660
DB	601	CACAGGTGATGCTGGAGCTGGCGGACGGGAAGCTCTTCCAAAGGAGATTTGATATCTC	660
QY	661	TTCTACATCACTGAGGAGATGAACAGAGTGCACGGATGAGCATGCAAGACCTC	720
DB	661	TTCTACATCACTGAGGAGATGAACAGAGTGCACGGATGAGCATGCAAGACCTC	720
QY	721	ATCTTCAGTGTGCTGAGCCCGAGCGGCTCTCCAGGAGCTCATCGAGTTCGCGAG	780
DB	721	ATCTTCAGTGTGCTGAGCCCGAGCGGCTCTCTCCAGGAGCTCATCGAGTTCGCGAG	780

QY	781	CGCTCTCTTTTCATCATCGACGGTTCGATGAGCTCAAGCCTTCTTTCCAGGATCCTCAG	840
DB	781	CGCTCTCTTTTCATCATCGACGGTTCGATGAGCTCAAGCCTTCTTTCCAGGATCCTCAG	840
QY	841	GGACCCCTGGTGCCTCTGCTGGAGGAGAAACGGCCACGAGCTGCTTCTTAACAGCTTA	900
DB	841	GGACCCCTGGTGCCTCTGCTGGAGGAGAAACGGCCACGAGCTGCTTCTTAACAGCTTA	900
QY	901	ATTTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACACGCGCCACGGCTTG	960
DB	901	ATTTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACACGCGCCACGGCTTG	960
QY	961	GAGAGCTCCACCGTCTGCTGGAGCACCCACGAGCTGAGAGTCTGGGGTCTCTTCAG	1020
DB	961	GAGAGCTCCACCGTCTGCTGGAGCACCCACGAGCTGAGAGTCTGGGGTCTCTTCAG	1020
QY	1021	GCAGAAAGGAGGAATATCTTCTACAAGTATTTCCACAATGAGAGAGGGGGCCAAAGTC	1080
DB	1021	GCAGAAAGGAGGAATATCTTCTACAAGTATTTCCACAATGAGAGAGGGGGCCAAAGTC	1080
QY	1081	TTCAATTACGTGAGGGAACAAGAGCTCTCTTTCACCATGCTGCTTCCCTCCCTGGTGTGC	1140
DB	1081	TTCAATTACGTGAGGGAACAAGAGCTCTCTTTCACCATGCTGCTTCCCTCCCTGGTGTGC	1140
QY	1141	TGGTGTGTGTACTCTGCTCCAGCAGCTGAGAGCTGGGGGCTGTTGAGACAGACG	1200
DB	1141	TGGTGTGTGTACTCTGCTCCAGCAGCTGAGAGCTGGGGGCTGTTGAGACAGACG	1200
QY	1201	TCCAGGACCACTGTCAGTGTATCATGCTCTTACCTGCTGAGCTCTGATGCAACCCAAAGCG	1260
DB	1201	TCCAGGACCACTGTCAGTGTATCATGCTCTTACCTGCTGAGCTCTGATGCAACCCAAAGCG	1260
QY	1261	GGGGCCCCGGCCCTCCAGCCCCCAACCAACAGAGAGGGTGTGCTCTTGGCGGCAAT	1320
DB	1261	GGGGCCCCGGCCCTCCAGCCCCCAACCAACAGAGAGGGTGTGCTCTTGGCGGCAAT	1320
QY	1321	GGGCTCTGGAAATCAGAAATCTTATTTGAGAGAGGAGCTCCGGAGACAGGCTTAGAC	1380
DB	1321	GGGCTCTGGAAATCAGAAATCTTATTTGAGAGAGGAGCTCCGGAGACAGGCTTAGAC	1380
QY	1381	GGGGAAGACGCTCTGCTTCTCAACATGAACATCTTCCAGAGAGGACATCACTGTGAG	1440
DB	1381	GGGGAAGACGCTCTGCTTCTCAACATGAACATCTTCCAGAGAGGACATCACTGTGAG	1440
QY	1441	AGGTACTACAGCTTCACTCACTTTCAGGAAATCTTTCAGAGCTATGCTATATATC	1500
DB	1441	AGGTACTACAGCTTCACTCACTTTCAGGAAATCTTTCAGAGCTATGCTATATATC	1500
QY	1501	CTGGACGAGGGAGGG	1560
DB	1501	CTGGACGAGGGAGGG	1560
QY	1561	TACGGCTTTCTGAAAGGAGCTTCTGCGCATCTCACAGCGCTTCTCTGTTTGGACTCCTG	1620
DB	1561	TACGGCTTTCTGAAAGGAGCTTCTGCGCATCTCACAGCGCTTCTCTGTTTGGACTCCTG	1620
QY	1621	AACGAGGAGACAGGAGCCACTCTGGAAGAGTCTCTGCTGGAAGGCTCTGCGCGCACATC	1680
DB	1621	AACGAGGAGACAGGAGCCACTCTGGAAGAGTCTCTGCTGGAAGGCTCTGCGCGCACATC	1680
QY	1681	AAGATGAGCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGCTCCACCCCTGCGAG	1740
DB	1681	AAGATGAGCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGCTCCACCCCTGCGAG	1740
QY	1741	CAGGGCTCTCTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGAGAGGAGTTCATCCAG	1800
DB	1741	CAGGGCTCTCTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGAGAGGAGTTCATCCAG	1800
QY	1801	CAGGCGCTGAGGCACTTCAGGAGTGTGCTGTCAGCAATTCCTCCAGATGAGGAC	1860
DB	1801	CAGGCGCTGAGGCACTTCAGGAGTGTGCTGTCAGCAATTCCTCCAGATGAGGAC	1860
QY	1861	ATGGCTCTCTGCTCTGCTGAGGCGCTGTCAGGAGCGCCCGCTGCTGACCTTGTATGGC	1920
DB	1861	ATGGCTCTCTGCTCTGCTGAGGCGCTGTCAGGAGCGCCCGCTGCTGACCTTGTATGGC	1920

[illegible]

Db	2941	AACAAAGCCCTAGGGGACACAGGTGTCGAGTCTGCTTTTGCAGAGCGGCTGAGCCATCTCTGGC	3000
QY	3001	TGCAAACTCCGAGTCTCTGGTATTATTTGGGATGGACCTGAATAAAATGACCCACAGT	3057
Db	3001	TGCAAACTCCGAGTCTCTGGTATTATTTGGGATGGACCTGAATAAAATGACCCACAGT	3057
RESULT 3			
AAAL44363			
ID	AAAL44363	standard; cDNA; 3186 BP.	
XX			
AC	AAAL44363;		
XX			
DT	31-OCT-2002	(first entry)	
XX			
DE	Human PYRIN-8	cDNA sequence #2.	
XX			
KW	Human; gene; ss; gene therapy; PYRIN; stress-related response;		
KW	apoptotic response; inflammatory response; inflammatory disorder;		
KW	immune system disorder; Crohn's disease; multiple sclerosis; cancer;		
KW	leukemia; autoimmune disorder; arthritis; neurological disease;		
KW	Alzheimer's disease; Parkinson's disease; chromosomal mapping;		
KW	tissue typing; forensic biology; predictive medicine; pharmacogenomics;		
KW	transcription profiling; PYRIN-8.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3186	
FT		/*tag= a	
FT		/product= "Human PYRIN-8 protein #2"	
XX			
PN	WO200261049-A2.		
XX			
PD	08-AUG-2002.		
XX			
PF	31-JAN-2002; 2002WO-US002967.		
XX			
PR	31-JAN-2001; 2001US-0265231P.		
PR	10-SEP-2001; 2001US-0318645P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
PA	(AMHP) WYETH.		
XX			
PI	Bertin J, Wang W, Blatcher M;		
XX			
DR	WPI; 2002-627477/67.		
DR	P-FSDB; AAO15590.		
XX			
PT	New PYRIN polypeptides and nucleic acids useful for modulating and		
PT	diagnosing stress-related, apoptotic and inflammatory responses, or for		
PT	treating inflammatory and immune system disorders, cancers, or		
PT	neurological diseases.		
XX			
PS	Claim 4; Fig 8; 167pp; English.		
XX			
CC	The invention comprises the amino acid and coding sequences of human		
CC	PYRIN proteins. The PYRIN protein and DNA sequences of the invention are		
CC	useful for modulating and diagnosing stress-related, apoptotic and		
CC	inflammatory responses. The PYRIN protein and DNA sequences are useful		
CC	for treating: inflammatory disorders and immune system disorders (e.g.		
CC	Crohn's disease, reactive arthritis, multiple sclerosis, contact		
CC	dermatitis, psoriasis, graft rejection, allergies, viral infections and		
CC	bacterial infections); cancer (e.g. leukaemia); autoimmune disorders		
CC	(e.g. systemic lupus erythematosus and arthritis); and neurological		
CC	diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN		
CC	protein and DNA sequences may also be used in screening assays, detection		
CC	assays (e.g. chromosomal mapping, tissue typing or forensic biology),		
CC	predictive medicine (e.g. diagnostic assays, clinical trials and		
CC	pharmacogenomics) and transcription profiling. The present DNA sequence		
CC	encodes a human PYRIN-8 protein		
XX			
SQ	Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other;		

Query Match		79.6%;	Score 2475;	DB 6;	Length 3186;		
Best Local Similarity		99.8%;	Pred. No. 0;				
Matches 2815;		Conservative	0;	Mismatches	2;	Indels	3; Gaps 1;
QY	289	GATCCCGAGAAACCTACAGGACATATGTCGCGAGGAAATCCGGCTCATGGAAGACCGC	348				
DB	370	GATCCCGAGAAACCTACAGGACATATGTCGCGAGGAAATCCGGCTCATGGAAGACCGC	429				
QY	349	AATGCGCGCTAGGGAATGTCTCAACCTCAGCACCGGTACACCGGCTCTGCTGGTG	408				
DB	430	AATGCGCGCTAGGGAATGTCTCAACCTCAGCACCGGTACACCGGCTCTGCTGGTG	489				
QY	409	AAGGAGCACTCAAAACCCATGCGAGTCCAGCAGCAGCTTCTGGACACAGGCGCGGACAC	468				
DB	490	AAGGAGCACTCAAAACCCATGCGAGTCCAGCAGCAGCTTCTGGACACAGGCGCGGACAC	549				
QY	469	GCAGAGACCTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC	528				
DB	550	GCAGAGACCTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC	609				
QY	529	GAGGAGCGCCCGAGCCACCGCACCGTGTGTCATGCAAGCGCGGAGGATAGGCAAG	588				
DB	610	GAGGAGCGCCCGAGCCACCGCACCGTGTGTCATGCAAGCGCGGAGGATAGGCAAG	669				
QY	589	TCCATGCTGGCACACAAAGGTGATGCTGGAATGCGGCGGACCGGAGCTCTTCCAGGCGAG	648				
DB	670	TCCATGCTGGCACACAAAGGTGATGCTGGAATGCGGCGGACCGGAGCTCTTCCAGGCGAG	729				
QY	649	TTTGATATCTCTTCTACATCACTGCGAGGAGATGAACAGAGTGCACAGGATGCAGC	708				
DB	730	TTTGATATCTCTTCTACATCACTGCGAGGAGATGAACAGAGTGCACAGGATGCAGC	789				
QY	709	ATGCAAGACCTCATCTTTCAGTGTGCTGAGCCAGCGGCTCTCCAGGAGCTCATC	768				
DB	790	ATGCAAGACCTCATCTTTCAGTGTGCTGAGCCAGCGGCTCTCCAGGAGCTCATC	849				
QY	769	CGAGTTCGCGAGCGCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTCTTCTTC	828				
DB	850	CGAGTTCGCGAGCGCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTCTTCTTC	909				
QY	829	CAGATCTCTCAGGACCGCTGTGCTCTGCTGGAGGAGAAACGGCCGACAGGAGCTGCTT	888				
DB	910	CAGATCTCTCAGGACCGCTGTGCTCTGCTGGAGGAGAAACGGCCGACAGGAGCTGCTT	969				
QY	889	CTTAAACAGCTTAATTCGGAAGAAGCTGCTCCCTCAGCTATCTTTGCTCATCAACACACGG	948				
DB	970	CTTAAACAGCTTAATTCGGAAGAAGCTGCTCCCTCAGCTATCTTTGCTCATCAACACACGG	1029				
QY	949	CCACGCGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGCGATGTGGAGATCCTG	1008				
DB	1030	CCACGCGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGCGATGTGGAGATCCTG	1089				
QY	1009	GGCTTCTCAGGAGAGAAAGGAAGATATCTTCAAGATATTCACAAATGCAGAGCAG	1068				
DB	1090	GGCTTCTCAGGAGAGAAAGGAAGATATCTTCAAGATATTCACAAATGCAGAGCAG	1149				
QY	1069	GCGGCGCAAGTCTTCAATTAAGTCAGGAGAACAGGCTCTTTCACCATGCTTCTGTC	1128				
DB	1150	GCGGCGCAAGTCTTCAATTAAGTCAGGAGAACAGGCTCTTTCACCATGCTTCTGTC	1209				
QY	1129	CCCTGCTGTGCTGGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGCTGGGGGCTG	1188				
DB	1210	CCCTGCTGTGCTGGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGCTGGGGGCTG	1269				
QY	1189	TTGAGACAGAGCTCAGAGACACCACTGAGTGTATGCTCTACCTGCTGAGTCTGATG	1248				
DB	1270	TTGAGACAGAGCTCAGAGACACCACTGAGTGTATGCTCTACCTGCTGAGTCTGATG	1329				
QY	1249	CAACCCAGCGGGGGCGCGCTCCAGCGCCCAACCAACAGAGAGGCTTGTGCTCC	1308				
DB	1330	CAACCCAGCGGGGGCGCGCTCCAGCGCCCAACCAACAGAGAGGCTTGTGCTCC	1389				

QY	1309	TTGGCGGAGATGGGCTCTGGAAATCAGAAATCCTATTATTGAGGAGCAGGACCTCCGGAAG	1368				
DB	1390	TTGGCGGAGATGGGCTCTGGAAATCAGAAATCCTATTATTGAGGAGCAGGACCTCCGGAAG	1449				
QY	1369	CACGGCTAGACGGGAGAGACGCTCTGCTCCTCAACATGAACATCTTCCAGAAAGGAC	1428				
DB	1450	CACGGCTAGACGGGAGAGACGCTCTGCTCCTCAACATGAACATCTTCCAGAAAGGAC	1509				
QY	1429	ATCAACTGTGAGAGGTACTACAGCTTCACTGAGTTTCCAGGAATTTCTTTGACGCT	1488				
DB	1510	ATCAACTGTGAGAGGTACTACAGCTTCACTGAGTTTCCAGGAATTTCTTTGACGCT	1569				
QY	1489	ATGTAATATCTTGACGAGGGGAGGGCGGAGCCAGACACAGGACGTCACAGG	1548				
DB	1570	ATGTAATATCTTGACGAGGGGAGGGCGGAGCCAGACACAGGACGTCACAGG	1629				
QY	1549	CTGTTGACCGAGTACGCTTTCTGAAAGGAGCTTCTGCACTCACACGCGCTTCTG	1608				
DB	1630	CTGTTGACCGAGTACGCTTTCTGAAAGGAGCTTCTGCACTCACACGCGCTTCTG	1689				
QY	1609	TTTGGACTCTCTGAACGAGGAGACACAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC	1668				
DB	1690	TTTGGACTCTCTGAACGAGGAGACACAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC	1749				
QY	1669	TCGCGGACATCAAGATGAGCTGTTGCACTGGATCCAAAGCAAGCTCAGAGCGCGC	1728				
DB	1750	TCGCGGACATCAAGATGAGCTGTTGCACTGGATCCAAAGCAAGCTCAGAGCGCGC	1809				
QY	1729	TCACACCTGACAGGAGCTCTTGGAGTCTCTCAGCTGCTTGTACGAGATCCAGAGGAG	1788				
DB	1810	TCACACCTGACAGGAGCTCTTGGAGTCTCTCAGCTGCTTGTACGAGATCCAGAGGAG	1869				
QY	1789	GAGTTATCCAGAGGCGCTTGAGCCACTTCCAGGTGATCGTGGTACGACCAATTCCTCC	1848				
DB	1870	GAGTTATCCAGAGGCGCTTGAGCCACTTCCAGGTGATCGTGGTACGACCAATTCCTCC	1929				
QY	1849	AGATGAGACATAGTCTCTCTGCTGCTGAGCGCTGACAGGCGCCAGGCTGCTG	1908				
DB	1930	AGATGAGACATAGTCTCTCTGCTGCTGAGCGCTGACAGGCGCCAGGCTGCTG	1989				
QY	1909	CACCTGTATGCGCCACTTACAGCGGAGCGGGAAGACCGCGAGGTGCTCCGCGAGGA	1968				
DB	1990	CACCTGTATGCGCCACTTACAGCGGAGCGGGAAGACCGCGAGGTGCTCCGCGAGGA	2049				
QY	1969	GCGCACACGCTGTTGTCAGCTCAGACAGAGAGGACCGTCTGCTGAGCGCTTACAGT	2028				
DB	2050	GCGCACACGCTGTTGTCAGCT---ACCAGAGAGGACCGTCTCTGCTGGACGCTACAGT	2106				
QY	2029	GAACATCTGGAGCGGCGCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA	2088				
DB	2107	GAACATCTGGAGCGGCGCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA	2166				
QY	2089	AATCCCTTGGGACGCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCAACTGC	2148				
DB	2167	AATCCCTTGGGACGCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCAACTGC	2226				
QY	2149	AAACTTTCAGAACCTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTCGAGAGACCTC	2208				
DB	2227	AAACTTTCAGAACCTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTCGAGAGACCTC	2286				
QY	2209	TCTCAGCTCTCATAGCAATGAAGTTGACAGAGTGAATCTCAGTGGCAACGGGCTT	2268				
DB	2287	TCTCAGCTCTCATAGCAATGAAGTTGACAGAGTGAATCTCAGTGGCAACGGGCTT	2346				
QY	2269	GGATTCCAGGATGATGCTCTTTTGGAGGGCGCTGGGCAATCCCAAGTCAGGCTGAG	2328				
DB	2347	GGATTCCAGGATGATGCTCTTTTGGAGGGCGCTGGGCAATCCCAAGTCAGGCTGAG	2406				
QY	2329	ATGATTTCAGTTGAGGAAGTGTACGCTGAGTTCGCGGGCTTGTGAGGAGATGGCTTCTGTG	2388				
DB	2407	ATGATTTCAGTTGAGGAAGTGTACGCTGAGTTCGCGGGCTTGTGAGGAGATGGCTTCTGTG	2466				
QY	2389	CTCGGACCAACCCCATCTCTGTTGAGTTGAGCTTGACCTGACAGGAATGCATCGAGGATTG	2448				

Db	678	CGAGGACCGGTGGGACACCAAGCTAGCGCCCATCAAGATAGAGACCCCTCTTTTGAGCCAGACG	737
Qy	530	AGGAGCCGCCGAGCCAGCCGCAACCGTGTGTCATGCAAGGCGCGGAGGATAGGCAAGT	589
Db	738	AGGAGCCGCCGAGCCAGCCGCAACCGTGTGTCATGCAAGGCGCGGAGGATAGGCAAGT	797
Qy	590	CCATGCTGGGACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAAGGCAGAT	649
Db	798	CCATGCTGGGACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAAGGCAGAT	857
Qy	650	TTGATTTATCTTCTTACATCAATGCAAGGAGATGAACAGAGTGCACGGATGACGCA	709
Db	858	TTGATTTATCTTCTTACATCAATGCAAGGAGATGAACAGAGTGCACGGATGACGCA	917
Qy	710	TGCAAGACCTCATCTTCAGCTGTGGCTTCAGCCGAGCGGCTCTTCAGGAGCTCATCC	769
Db	918	TGCAAGACCTCATCTTCAGCTGTGGCTTCAGCCGAGCGGCTCTTCAGGAGCTCATCC	977
Qy	770	GAGTTCCGAGCGCCTCTTTTCATCATGACGGCTTCGATGAGCTCAAGCCTCTTTTCC	829
Db	978	GAGTTCCGAGCGCCTCTTTTCATCATGACGGCTTCGATGAGCTCAAGCCTCTTTTCC	1037
Qy	830	ACGATCTCTCAGGACCTCTGCTGCTGGGAGGAGAAACGGCCCAACGGAGCTCTTC	889
Db	1038	ACGATCTCTCAGGACCTCTGCTGCTGGGAGGAGAAACGGCCCAACGGAGCTCTTC	1097
Qy	890	TTAACAGCTTAATTCGGAAGAAGCTGCTCCTGAGCTATCTTTTGCTCATCACACACGGC	949
Db	1098	TTAACAGCTTAATTCGGAAGAAGCTGCTCCTGAGCTATCTTTTGCTCATCACACACGGC	1157
Qy	950	CCACGGCTTTGGAGAAGCTCCACGCTGCTGGAGGACCCACGGGCATGTGGAGATCCTGG	1009
Db	1158	CCACGGCTTTGGAGAAGCTCCACGCTGCTGGAGGACCCACGGGCATGTGGAGATCCTGG	1217
Qy	1010	GCTTCTCTGAGGACGAAGAAGAAATCTTCAAGATATTTCCAAATGCAAGACGAGG	1069
Db	1218	GCTTCTCTGAGGACGAAGAAGAAATCTTCAAGATATTTCCAAATGCAAGACGAGG	1277
Qy	1070	CGGSCCAAGTCTTCAAATTTAGTGAGGACCAAGAGCCTCTTCCACCATGTCTCGTCC	1129
Db	1278	CGGSCCAAGTCTTCAAATTTAGTGAGGACCAAGAGCCTCTTCCACCATGTCTCGTCC	1337
Qy	1130	CCCTGSGTGTCTGGGTGSGTGTAACCTGCTCCAGCAGCAGCTGGAGGTTGGGGGCTGT	1189
Db	1338	CCCTGSGTGTCTGGGTGSGTGTAACCTGCTCCAGCAGCAGCTGGAGGTTGGGGGCTGT	1397
Qy	1190	TGAGACAGACGTCCAGGACACCACTGCGAGTGTACGTCTTACCTGCTGAGTCTGATGC	1249
Db	1398	TGAGACAGACGTCCAGGACACCACTGCGAGTGTACGTCTTACCTGCTGAGTCTGATGC	1457
Qy	1250	AACCCAGCCGGGGGCCCGCGCTCCAGCCCCCACCACCAACAGAGAGGTTGTGCTCT	1309
Db	1458	AACCCAGCCGGGGGCCCGCGCTCCAGCCCCCACCACCAACAGAGAGGTTGTGCTCT	1517
Qy	1310	TGGCGGCAGATGGGCTCTGGAATCAGAAAATCTTATTTGAGGACGAGCACTCCGGAAGC	1369
Db	1518	TGGCGGCAGATGGGCTCTGGAATCAGAAAATCTTATTTGAGGACGAGCACTCCGGAAGC	1577
Qy	1370	ACGGCCTAGACGGGGAAGACGTCTCTGCTTCTCAAATGAAATCTTCCAGAGGACA	1429
Db	1578	ACGGCCTAGACGGGGAAGACGTCTCTGCTTCTCAAATGAAATCTTCCAGAGGACA	1637
Qy	1430	TCAACTGTGAGAGTACTACAGCTTTCATCCACTTTGAGTTTCCAGGAATCTTTGCAGCTA	1489
Db	1638	TCAACTGTGAGAGTACTACAGCTTTCATCCACTTTGAGTTTCCAGGAATCTTTGCAGCTA	1697
Qy	1490	TGTAATAATCTCTGGAACGAGGGGAGGGCGGGGAGGCCACAGACGAGCACTGACCAAGC	1549
Db	1698	TGTAATAATCTCTGGAACGAGGGGAGGGCGGGGAGGCCACAGACGAGCACTGACCAAGC	1757
Qy	1550	TGTTGACCGAGTACCGGTTTCTGAAAGGAGCTTCTTGGAACCTCACAGCGGCTTCTGT	1609
Db	1758	TGTTGACCGAGTACCGGTTTCTGAAAGGAGCTTCTTGGAACCTCACAGCGGCTTCTGT	1817

RESULT 5

ABS/78719
ID ABS78719 standard; cDNA; 2158 BP.
AC ABS78719;
XX
XX 16-DEC-2002 (first entry)
XX
XX Human cDNA encoding NAAP9, from INCYTE no.429930CB1.
XX
XX Human; ss; gene; nucleic acid associated protein; NAAP; cancer;
KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.
XX
OS Homo sapiens.
XX
XX WO200272630-A2.
XX
XX 19-SEP-2002.
XX
XX 07-FEB-2002; 2002WO-US003844.
XX
XX 09-FEB-2001; 2001US-0269118P.
XX 21-FEB-2001; 2001US-0270963P.
XX 22-FEB-2001; 2001US-0270858P.
XX 23-FEB-2001; 2001US-0271194P.
XX 07-MAR-2001; 2001US-0274071P.
XX 12-APR-2001; 2001US-0283496P.
XX 09-NOV-2001; 2001US-0344650P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Malia NK, Thangavelu K;
PI Elliott VS, Marquis JP;
XX
XX WPI: 2002-723320/78.
XX P-PSDB; ABG97475.
XX
XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PI treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX
XX Claim 5; Page 160; 162pp; English.
XX
XX The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
CC screening for a compound that is effective as an ant/agonist or modulator
CC of NAAP, generating an expression profile of a sample containing the
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridisable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired

CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
CC other diseases and disorders listed in the specification. These are also
CC useful in assessing the effects of exogenous compounds on the expression
CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The microarray is useful in monitoring
CC or measuring protein-protein interactions, drug-target interactions, and
CC gene expression profiles. The present sequence encodes an NAAP protein
XX
SQ Sequence 2158 BP; 489 A; 578 C; 639 G; 452 T; 0 U; 0 Other;
Query Match 43.6%; Score 1356; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 1408 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 1467
DB 1 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 60
QY 1468 TTCAGGAAATCTTTGACGCTATGTACTATATCTTGGACGAGGGGGGGGAGGC 1527
DB 61 TTCAGGAAATCTTTGACGCTATGTACTATATCTTGGACGAGGGGGGGGAGGC 120
QY 1528 CCAGACGAGGACGTCACCGGCTGTGACCGAGTACGCGTTTCTGAAAGAGCTTCCTG 1587
DB 121 CCAGACGAGGACGTCACCGGCTGTGACCGAGTACGCGTTTCTGAAAGAGCTTCCTG 180
QY 1588 GCATCTCACCGCGCTTCTCTGTTGGACTCTCTGAACAGGAGGACAGGAGCCACTGGAG 1647
DB 181 GCATCTCACCGCGCTTCTCTGTTGGACTCTCTGAACAGGAGGACAGGAGCCACTGGAG 240
QY 1648 AAGAGTCTCTGCTGGAAAGGTCTGCGCGCACATCAAGATGGACCTTTGACGTGGATCAA 1707
DB 241 AAGAGTCTCTGCTGGAAAGGTCTGCGCGCACATCAAGATGGACCTTTGACGTGGATCAA 300
QY 1708 AGCAAGCTCAGAGGAGCGGCTCCACCTGACGAGGGGCTCTTGAGTCTTTCAGTGC 1767
DB 301 AGCAAGCTCAGAGGAGCGGCTCCACCTGACGAGGGGCTCTTGAGTCTTTCAGTGC 360
QY 1768 TTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCACTTCCAGGTGATC 1827
DB 361 TTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCACTTCCAGGTGATC 420
QY 1828 GTGTGTCAGCAATTCGCTCCAAAGATGGAGACATGTTCTCTGTTCTGTCTGAAGCGC 1887
DB 421 GTGTGTCAGCAATTCGCTCCAAAGATGGAGACATGTTCTCTGTTCTGTCTGAAGCGC 480
QY 1888 TGCAGGAGCGCCAGGTGCTGCACTTGTATGGCGCCACTTACAGCGCGGAGCGGGAAGAC 1947
DB 481 TGCAGGAGCGCCAGGTGCTGCACTTGTATGGCGCCACTTACAGCGCGGAGCGGGAAGAC 540
QY 1948 CGCGGAGGTGCTCCGAGGAGCGCACACGCTGTTGGTGACGCTCAGACCCAGAGAGGACC 2007
DB 541 CGCGGAGGTGCTCCGAGGAGCGCACACGCTGTTGGTGACGCT---ACCAGAGAGGACC 597
QY 2008 GTTCTGTGGACGCTACAGTGAACATCTGCGAGCGGCCCTGTGACCAATCCAAACCTG 2067
DB 598 GTTCTGTGGACGCTACAGTGAACATCTGCGAGCGGCCCTGTGACCAATCCAAACCTG 657
QY 2068 ATAGAGTGTCTGTATCCGAAATTCCTGGGAGCGGGGGGTGAAGCTGCTCTGTCAA 2127
DB 658 ATAGAGTGTCTGTATCCGAAATTCCTGGGAGCGGGGGGTGAAGCTGCTCTGTCAA 717
QY 2128 GGACTTCAGACACCCCACTGCAAACTTCAGAACCTGAGCGCTGAAGAGGTGCGGATCTCC 2187
DB 718 GGACTTCAGACACCCCACTGCAAACTTCAGAACCTGAGCGCTGAAGAGGTGCGGATCTCC 777
QY 2188 AGCTCAGCTCGAGGAGCTCTCTGCGAGCTCTCATAGCAATAAGAAATTTGACAGGATG 2247
DB 778 AGCTCAGCTCGAGGAGCTCTCTGCGAGCTCTCATAGCAATAAGAAATTTGACAGGATG 837

QY 2248 GATCTCAGTGGCAACGGCGTTGGAATCCAGGAGATGATGCTGCTTTCGAGGCGCTGCGG 2307
Dd |||||||
838 GATCTCAGTGGCAACGGCGTTGGAATCCAGGAGATGATGCTGCTTTCGAGGCGCTGCGG 897
QY 2308 CATCCCCAGTGGCGCTGAGATGATTCAGTTGAGGAGTGTGAGTGGAGTGGGGCT 2367
Dd |||||||
898 CATCCCCAATGCGAGCTGCGAGATGATTCAGTTGAGGAGTGTGAGTGGAGTGGGGCT 957
QY 2368 TGTGAGGAGATGGCTTCTGCTGCTGCGGACCAACCAATCTGCTGAGTTGGACCTGACA 2427
Dd |||||||
958 TGTGAGGAGATGGCTTCTGCTGCTGCGGACCAACCAATCTGCTGAGTTGGACCTGACA 1017
QY 2428 GGAATGCACTGGAGGATTTGGGCTGAGTTACTATGCGAGGAGTGGAGCAACCCAGTC 2487
Dd |||||||
1018 GGAATGCACTGGAGGATTTGGGCTGAGTTACTATGCGAGGAGTGGAGCAACCCAGTC 1077
QY 2488 TGCAGACTACGAGTCTGCTGCTGAGATCTGCGGCTCACTGCTGCTGCTGCTGAGGAG 2547
Dd |||||||
1078 TGCAGACTACGAGTCTGCTGCTGAGATCTGCGGCTCACTGCTGCTGCTGCTGAGGAG 1137
QY 2548 CTGGCTCACTCTCAGTGTGAACAGAGCTGAGAGAGTGGACCTGAGCCTGAATGAG 2607
Dd |||||||
1138 CTGGCTCACTCTCAGTGTGAACAGAGCTGAGAGAGTGGACCTGAGCCTGAATGAG 1197
QY 2608 CTGGGAGCTCGGGGTGCTGCTGCTGTGTGAGGGCTTCAGGCATCCAGTGCAGAGCTC 2667
Dd |||||||
1198 CTGGGAGCTCGGGGTGCTGCTGCTGTGTGAGGGCTTCAGGCATCCAGTGCAGAGCTC 1257
QY 2668 CAGACCTCGGTTGGGCTGCTGCGGCTGGCTGCTGCGGCTGCTGAGGCTCTTCTGTG 2727
Dd |||||||
1258 CAGACCTCGGTTGGGCTGCTGCGGCTGGCTGCTGCGGCTGCTGAGGCTCTTCTGTG 1317
QY 2728 GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTGAAGTTCAACGAGCTGGGAGAC 2787
Dd |||||||
1318 GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTGAAGTTCAACGAGCTGGGAGAC 1377
QY 2788 TGGGGCTGTGTTGTGCTGAGGGCTCAACATCCCGCTGCGAGACTCCAGAAACTG 2847
Dd |||||||
1378 TGGGGCTGTGTTGTGCTGAGGGCTCAACATCCCGCTGCGAGACTCCAGAAACTG 1437
QY 2848 TGGCTGGATGCTGTGGCTCAGACCAAGCTTGTGAGATCTTTACTTTCACCTGGG 2907
Dd |||||||
1438 TGGCTGGATGCTGTGGCTCAGACCAAGCTTGTGAGATCTTTACTTTCACCTGGG 1497
QY 2908 ATCAACAGACCTTGACCGGCTTTTACCTGACCAACAGCCCTAGGGGACACAGGTGTC 2967
Dd |||||||
1498 ATCAACAGACCTTGACCGGCTTTTACCTGACCAACAGCCCTAGGGGACACAGGTGTC 1557
QY 2968 CGACTGCTTTCAGCGGCTGAGCCATCTGCTGCTGCAAACTCCGAGTCTCTGTTATTT 3027
Dd |||||||
1558 CGACTGCTTTCAGCGGCTGAGCCATCTGCTGCTGCAAACTCCGAGTCTCTGTTATTT 1617
QY 3028 GGGATGACCTGAATAAATGACCCAGTAGGTTGGCAGGCTTCGAGTAACAAACCT 3087
Dd |||||||
1618 GGGATGACCTGAATAAATGACCCAGTAGGTTGGCAGGCTTCGAGTAACAAACCT 1677
QY 3088 TATTGGACATTTGGCTGCTCA 3108
Dd |||||||
1678 TATTGGACATTTGGCTGCTCA 1698

RESULT 6
ACD03624

ID ACD03624 standard; cDNA; 1683 BP.

XX ACD03624;

XX AC

XX AC

DT 01-AUG-2003 (first entry)

XX Novel human GPCR related protein NOV2b cDNA.

XX Human; G-protein coupled receptor related protein; GPCR related protein;

KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
KW diabetes; immune disorder; AIDS; obesity; asthma;
KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
KW infection; multiple sclerosis; cancer-associated cachexia;
KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;
KW gene; ss.

XX Homo sapiens.

OS XX

PN WO200299116-A2.

XX 12-DEC-2002.

PD 12-DEC-2002.

XX 04-JUN-2002; 2002WO-US017428.

XX 04-JUN-2001; 2001US-0295607P.

PR 04-JUN-2001; 2001US-0295661P.

PR 06-JUN-2001; 2001US-0296404P.

PR 06-JUN-2001; 2001US-0296418P.

PR 14-JUN-2001; 2001US-0298285P.

PR 15-JUN-2001; 2001US-0298556P.

PR 21-JUN-2001; 2001US-0299949P.

PR 26-JUN-2001; 2001US-0300883P.

PR 28-JUN-2001; 2001US-0301550P.

PR 13-AUG-2001; 2001US-0311372P.

PR 27-AUG-2001; 2001US-0315071P.

PR 29-AUG-2001; 2001US-0315660P.

PR 14-SEP-2001; 2001US-0322293P.

PR 17-SEP-2001; 2001US-0322706P.

PR 14-DEC-2001; 2001US-0341186P.

PR 28-FEB-2002; 2002US-0361189P.

PR 12-MAR-2002; 2002US-0363673P.

PR 12-MAR-2002; 2002US-0363676P.

PR 03-JUN-2002; 2002US-00363676.

XX (CURA-) CURAGEN CORP.

PA Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;

XX Gangoli EA, Gerlach VL, Gorman L, Guo X, Hjalte T, Kekuda R, Li L;

PI MacDougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;

PI Pena CE, Rastelli L, Shimkets RA, Stone DJ, Spytek KA, Vernet CAM;

PI Voss EZ, Zerhusen BD;

XX WPI; 2003-140627/13.

DR P-PSDB; ABU99120.

XX New NOVX polypeptides and nucleic acids, useful for preventing or

DR treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, or

XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

PT Claim 20; Page 99-100; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27

CC 118-961 residue amino acid sequences, given in the specification, a

CC mature form of them, a sequence that is at least 95 % identical to them,

CC or a sequence having one or more conservative substitutions in them. The

CC polypeptide is useful in manufacturing a medicament for treating a

CC syndrome associated with a human disease selected from a pathology

CC associated with the polypeptide. The NOVX polypeptides, polynucleotides

CC and antibodies are useful in treating or preventing NOVX-associated

CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune

CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's

CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-

CC associated cachexia, and other wasting disorders associated with chronic

CC diseases. The nucleic acids and polypeptides may also be used as targets

CC for the identification of small molecules that modulate or inhibit e.g.

CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,

CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This sequence encodes a novel human G-protein coupled receptor
CC related protein NOV

XX
SQ Sequence 1683 BP; 343 A; 483 C; 508 G; 349 T; 0 U; 0 Other;

Query Match 39.0%; Score 1212; DB 7; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1104	GCCTCTCTTACCATGCTCTTCTGCTCCCTCGTGTGCTGGTGTGTCTACTGCTCTCA	1163
DB	5	GCCTCTCTTACCATGCTCTTCTGCTCCCTCGTGTGCTGGTGTGTCTACTGCTCTCA	64
QY	1164	GCAGCAGCTGGAGGGTGGGGGCTGTTTGAGACAGACGTCCAGGACCACTGTCAGTGA	1223
DB	65	GCAGCAGCTGGAGGGTGGGGGCTGTTTGAGACAGACGTCCAGGACCACTGTCAGTGA	124
QY	1224	CATGCTCTACCTGTGAGTCTGATGCAACCAAGCCGGGGGCCCGCCCTCCAGCCCC	1283
DB	125	CATGCTCTACCTGTGAGTCTGATGCAACCAAGCCGGGGGCCCGCCCTCCAGCCCC	184
QY	1284	ACCCAAACAGAGAGGGTGTGCTCTCGCGCAGATGGGCTCTGGAATCAGAAATCCT	1343
DB	185	ACCCAAACAGAGAGGGTGTGCTCTCGCGCAGATGGGCTCTGGAATCAGAAATCCT	244
QY	1344	ATTTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGACGGGAAGACGCTCTCTGCCCT	1403
DB	245	ATTTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGACGGGAAGACGCTCTCTGCCCT	304
QY	1404	CAACATGAACATCTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCACTT	1463
DB	305	CAACATGAACATCTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCACTT	364
QY	1464	GAGTTTCAGAAATCTTTGAGCTATCTACTATATCTGACGAGGGGGAGGGGGG	1523
DB	365	GAGTTTCAGAAATCTTTGAGCTATCTACTATATCTGACGAGGGGGAGGGGGG	424
QY	1524	AGGCCACAGCAGGACGTGACAGGCTGTTGACCGAGTACGCGTCTTCTGAAAGAGCTT	1583
DB	425	AGGCCACAGCAGGACGTGACAGGCTGTTGACCGAGTACGCGTCTTCTGAAAGAGCTT	484
QY	1584	CCTGGCACTACAGCCGCTCTGTTGGACTCTGAAAGGAGAGACAGAGCCACT	1643
DB	485	CCTGGCACTACAGCCGCTCTGTTGGACTCTGAAAGGAGAGACAGAGCCACT	544
QY	1644	GGAGAAGAGTCTCTGGAAGGCTCGCCGACATCAAGATGACCTGTTGCACTGGAT	1703
DB	545	GGAGAAGAGTCTCTGGAAGGCTCGCCGACATCAAGATGACCTGTTGCACTGGAT	604
QY	1704	CCAAAGCAAAAGCTCAGAGCGAGGCTCCACCCTGTCAGCAGGCGCTCTTGGAGTCTTCAG	1763
DB	605	CCAAAGCAAAAGCTCAGAGCGAGGCTCCACCCTGTCAGCAGGCGCTCTTGGAGTCTTCAG	664
QY	1764	CTGCTTGTACAGATCCAGAGGAGGAGTATTCAGCAGGCGCTGAGCCACTCCAGGT	1823
DB	665	CTGCTTGTACAGATCCAGAGGAGGAGTATTCAGCAGGCGCTGAGCCACTCCAGGT	724
QY	1824	GATCGTGTGTCAGCAATTCCTCCAGATGAGACATGCTCTCTGTTCTGCTCA	1883
DB	725	GATCGTGTGTCAGCAATTCCTCCAGATGAGACATGCTCTCTGTTCTGCTCA	784
QY	1884	GGCTGTCAGGAGCGCCAGGTGCTGCATCTGATGCGGCCACTACAGCGGAGCGGGA	1943
DB	785	GGCTGTCAGGAGCGCCAGGTGCTGCATCTGATGCGGCCACTACAGCGGAGCGGGA	844
QY	1944	AGACCGCGAGGTGCTCCAGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACAGAG	2003
DB	845	AGACCGCGAGGTGCTCCAGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACAGAG	904

RESULT 7

ACD03623

ID ACD03623 standard; cDNA; 1800 BP.

XX AC ACD03623;

XX AC ACD03623;

DT 01-AUG-2003 (first entry)

XX

DE Novel human GPCR related protein NOV2a cDNA.

XX

Human; G-protein coupled receptor related protein; GPCR related protein;
NOV; cytosolic; cardiant; antiarteriosclerotic; antidiabetic;
immunomodulator; anti-HIV; anorectic; antiaesthetic; haemostatic;
antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
diabetes; immune disorder; AIDS; obesity; asthma;
haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
infection; multiple sclerosis; cancer-associated cachexia;
wasting disorder; chronic disease; neurogenesis; cell differentiation;
cell proliferation; haematopoiesis; wound healing; angiogenesis;
chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;
gene; ss.

OS Homo sapiens.

XX

XX WO200299116-A2.

XX

PD 12-DEC-2002.

XX

XX 04-JUN-2002; 2002WO-US017428.

XX

PR 04-JUN-2001; 2001US-0295607P.

PR 04-JUN-2001; 2001US-0295607P.

PR 06-JUN-2001; 2001US-0296404P.

PR 06-JUN-2001; 2001US-0296404P.

PR 14-JUN-2001; 2001US-0298285P.

PR 15-JUN-2001; 2001US-0298566P.

PR 21-JUN-2001; 2001US-0299949P.

PR 26-JUN-2001; 2001US-0300883P.

PR 28-JUN-2001; 2001US-0301505P.

PR 13-AUG-2001; 2001US-0311972P.

PR 27-AUG-2001; 2001US-0315071P.

PR 29-AUG-2001; 2001US-0315660P.

PR 14-SEP-2001; 2001US-0322293P.

PR 17-SEP-2001; 2001US-0322706P.

PR 14-DEC-2001; 2001US-0341186P.

PR 28-FEB-2002; 2002US-0361189P.

RESULT 8

ABL59333
ID ABL59333 standard; cDNA; 4931 BP.

XX AC ABL59333;
XX DT 07-OCT-2002 (first entry)

Nucleotide sequence of human leucine-rich repeat protein HLRRBM1.

XX Human; leucine-rich repeat; HLRRBM1; proliferative disorder;
XX immune condition; apoptosis; signal transduction; autoimmune disease;
XX haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
XX cardiovascular disorder; neurological disease; pheromone;
XX pulmonary disease; chronic obstructive pulmonary disease;
XX allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
XX haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
XX inflammatory disorder; systemic lupus erythematosus;
XX cardiovascular disease; cancer; gene; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1320..2666
XX /*tag= a
XX /*product= "HLRRBM1"

XX WO200252011-A2.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-US049740.

XX 22-DEC-2000; 2000US-0257773P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder J, Ramanathan C, Mintier G;

XX WPI; 2002-566676/60.

XX P-PSDB; ABB77910.

XX New HLRRBM1 nucleic acids for preventing, treating or ameliorating e.g.
XX proliferative disorders, immune conditions, a disorder related to
XX aberrant apoptosis modulation or developmental disorders.

XX Claim 1; Fig 1A-E; 371pp; English.

XX The present sequence encodes a human leucine-rich repeat containing
XX protein, designated HLRRBM1. HLRRBM1 polypeptides and polynucleotides are
XX useful for preventing, treating or ameliorating a medical condition such
XX as a proliferative disorder, immune condition, or a disorder related to
XX aberrant apoptosis modulation, either directly or indirectly, and in
XX modulating signal transduction activity in various cells, tissue and
XX organisms. They are also useful for treating, preventing, or diagnosing
XX diseases of haematopoietic cells, autoimmune disease, graft-versus-host
XX disease, allergic conditions (e.g. asthma), cardiovascular disorders, and
XX neurological diseases, and for increasing the organisms' ability to
XX synthesize and/or release pheromones. The polypeptide may also be used in
XX treating, preventing or ameliorating pulmonary disease (e.g. chronic
XX obstructive pulmonary disease, allergic rhinitis, or bronchial
XX hyperresponsiveness), reproductive disease, haematopoietic disease,
XX platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious
XX disorders (e.g. innate immunity to bacterial pathogens, or adaptive
XX immune response), immune and inflammatory disorders (e.g. systemic lupus
XX erythematosus), cardiovascular diseases and cancers. HLRRBM1 nucleic
XX acids may further be used in chromosome identification or mapping, as a
XX chromosome marker, as molecular weight markers, as diagnostic probes, in
XX gene therapy, in raising anti-DNA antibodies, or as antigens for
XX eliciting immune responses

SQ Sequence 4931 BP; 1182 A; 1357 C; 1331 G; 1061 T; 0 U; 0 Other;
Query Match 24.7%; Score 768; DB 6; Length 4931;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	676	AGGAGATGAACACAGAGTGCACGGAATGAGCATGCAAGACCTCATCTTCAGTGTGG	735
DB	1314	AGGAGATGAACACAGAGTGCACGGAATGAGCATGCAAGACCTCATCTTCAGTGTGG	1373
QY	736	CTGTAGCCAGCGGCTCTCCAGAGCTCATCGAGTCCGAGCGCTCTTTTCATC	795
DB	1374	CTGTAGCCAGCGGCTCTCCAGAGCTCATCGAGTCCGAGCGCTCTTTTCATC	1433
QY	796	ATCGAGCGCTTCGATGAGCTCAAGCCTTCTTCCACGATCTCAGGGACCTCGTGCTC	855
DB	1434	ATCGAGCGCTTCGATGAGCTCAAGCCTTCTTCCACGATCTCAGGGACCTCGTGCTC	1493
QY	856	TGCTGGGAGGAGAAACCGGCCACGAGCTGCTTTTAAACAGCTTAATTGCGAAGAGCTG	915
DB	1494	TGCTGGGAGGAGAAACCGGCCACGAGCTGCTTTTAAACAGCTTAATTGCGAAGAGCTG	1553
QY	916	CTCCCTGAGCTATCTTTGCTCATCACACAGCGGCCACGGCTTTGGAGAGCTCCACCT	975
DB	1554	CTCCCTGAGCTATCTTTGCTCATCACACAGCGGCCACGGCTTTGGAGAGCTCCACCT	1613
QY	976	CTGCTGGAGACACCCAGGCAATGTGAGATCTCTGGGCTTCTCTGAGGCAGAAAGGAA	1035
DB	1614	CTGCTGGAGACACCCAGGCAATGTGAGATCTCTGGGCTTCTCTGAGGCAGAAAGGAA	1673
QY	1036	TACTTCTAAGATATTTCCAAATGCAGAGCAGCGGGCCAAAGTCTTCAATTACGTAGG	1095
DB	1674	TACTTCTAAGATATTTCCAAATGCAGAGCAGCGGGCCAAAGTCTTCAATTACGTAGG	1733
QY	1096	GACAAAGAGCCTCTCTTCCACCATGTCTTCCCTCCCTGCTGCTGGTGTGTGTACC	1155
DB	1734	GACAAAGAGCCTCTCTTCCACCATGTCTTCCCTCCCTGCTGCTGGTGTGTGTACC	1793
QY	1156	TGCTTCCAGCAGCAGCTGGAGGCTGGGGGGCTGTGTAGACAGAGCTCCAGGACCACT	1215
DB	1794	TGCTTCCAGCAGCAGCTGGAGGCTGGGGGGCTGTGTAGACAGAGCTCCAGGACCACT	1853
QY	1216	GCAGTGATACGTCTCTACCTCTGAGTCTGATGCAACCAAGCCGGGGCCCCGGCCTC	1275
DB	1854	GCAGTGATACGTCTCTACCTCTGAGTCTGATGCAACCAAGCCGGGGCCCCGGCCTC	1913
QY	1276	CAGCCCCCACCACCAACAGAGGGTGTGCTTCCCTGGCGGAGATGGGCTCTTGAATCAG	1335
DB	1914	CAGCCCCCACCACCAACAGAGGGTGTGCTTCCCTGGCGGAGATGGGCTCTTGAATCAG	1973
QY	1336	AAATCTCTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTTAGACGGGGAAGACCTCTCT	1395
DB	1974	AAATCTCTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTTAGACGGGGAAGACCTCTCT	2033
QY	1396	GCCTTCTCAATGAACATCTTCCAGAAAGGACATCAACTGTGAGAGG	1443
DB	2034	GCCTTCTCAATGAACATCTTCCAGAAAGGACATCAACTGTGAGAGG	2081

RESULT 9

ADC32201

ID ADC32201 standard; cDNA; 591 BP.

XX AC ADC32201;

XX DT 18-DEC-2003 (first entry)

XX Human novel cDNA contig sequence, SEQ ID NO:2283.

XX Human; diagnostic; drug screening; forensics; gene mapping;

XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;

XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

XX ulcers; osteoporosis; autoimmune disease; cancer;

PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 23437; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX SQ Sequence 487 BP; 97 A; 124 C; 155 G; 105 T; 0 U; 6 Other;

Query Match 12.4%; Score 385; DB 8; Length 487;
 Best Local Similarity 99.8%; Pred. No. 5.6e-165;
 Matches 435; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2212 GCAGCTCTCATAGCAATGAAGATTTGACAGGATGATCTCAGTGGCAACGGCGTTGGA 2271
 Db 31 GCAGCTCTCATAGCAATGAAGATTTGACAGGATGATCTCAGTGGCAACGGCGTTGGA 90

Qy 2272 TTCCAGCATGATGCTCTTTGGAGGGCTGGGGCTGCCAGTCCCGAGTCAGGCTGCAGATG 2331
 Db 91 TTCCAGCATGATGCTCTTTGGAGGGCTGGGGCTGCCAGTCCCGAGTCAGGCTGCAGATG 150

Qy 2332 ATTGAGTTGAGGAGTGTCAGCTGAGTCCGGGGCTGTCAGGAGATGGCTTCTGTGCTC 2391
 Db 151 ATTGAGTTGAGGAGTGTCAGCTGAGTCCGGGGCTGTCAGGAGATGGCTTCTGTGCTC 210

Qy 2392 GGCACCAACCCACATCTGGTTGAGTTGACCTGACAGGAAATGCACTGGAGATTTGGGC 2451
 Db 211 GGCACCAACCCACATCTGGTTGAGTTGACCTGACAGGAAATGCACTGGAGATTTGGGC 270

Qy 2452 CTGAGTTACTATGCCAGGACTGAGGACCCAGTCTGACACTACGAGCTTTGGGCTG 2511
 Db 271 CTGAGTTACTATGCCAGGACTGAGGACCCAGTCTGACACTACGAGCTTTGGGCTG 330

Qy 2512 AAGATCTCCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2571
 Db 331 AAGATCTCCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390

Qy 2572 CAGAGCTGAGAGAGCTGGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2631
 Db 391 CAGAGCTGAGAGAGCTGGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 450

Qy 2632 CTGTGTGAGGGCTTCA 2647
 Db 451 CTGTGTGAGGGCTTCA 466

RESULT 11
 AAC76566
 ID AAC76566 standard; cDNA; 479 BP.

XX AAC76566;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF2121 polynucleotide sequence SEQ ID NO:4241.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO2000058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42357.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 3436; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX Sequence 479 BP; 101 A; 139 C; 145 G; 93 T; 0 U; 1 Other;

Query Match 11.1%; Score 346; DB 3; Length 479;

Best Local Similarity 99.7%; Pred. No. 3.4e-147;

Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1562 ACGCGTTTCTGAAAGGAGCTCTCGGCACTCACCAGCGGCTTCTCTGTTGGACTCTGA 1621

Db 2 ACGCGTTTCTGAAAGGAGCTCTCGGCACTCACCAGCGGCTTCTCTGTTGGACTCTGA 61

Qy 1622 ACGAGGACACGAGGACCCAGGAGAGTCTCTGTTGGAAGGTCTCGCGGACATCA 1681

CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting an
 CC polynucleotide or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences.

XX Sequence 3306 BP; 718 A; 958 C; 939 G; 691 T; 0 U; 0 Other;

Query Match 9.3%; Score 289; DB 9; Length 3306;
 Best Local Similarity 100.0%; Pred. No. 3e-121;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGGACGGCCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Db 497 ATGCTACGAACCGCAGGACGGCCTCTGTGCGCTGTCCCTACTTGGAGAACTC 438
 Qy 61 GAGCGTGTGGAACGAAAGTTCAAGTTATACCTGGGACCGGACAGAGCTGGAGAA 120
 Db 437 GAGCGTGTGGAACGAAAGTTCAAGTTATACCTGGGACCGGACAGAGCTGGAGAA 378
 Qy 121 GCGAAGATCCCTCGGGAAGCATGGAAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Db 377 GCGAAGATCCCTCGGGAAGCATGGAAGGCGGCTCCCTGGAATGGCCAGCTGCTC 318
 Qy 181 ATCACCCTTGGGCGCAGAGAGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 240
 Db 317 ATCACCCTTGGGCGCAGAGAGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 258
 Qy 241 AACAGGAGGACCTTGGGAGAGGACAGAGAGGACCTGGTGAGGG 289
 Db 257 AACAGGAGGACCTTGGGAGAGGACAGAGAGGACCTGGTGAGGG 209

RESULT 14

AAK50424
 ID AAK50424 standard; DNA; 282 BP.

XX AC AAK50424;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 24981.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US0000668.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 DR Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 PS Example 4; SEQ ID NO 24981; 658pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention

SQ Sequence 282 BP; 71 A; 71 C; 96 G; 44 T; 0 U; 0 Other;

Query Match 9.1%; Score 282; DB 4; Length 282;
 Best Local Similarity 100.0%; Pred. No. 5.4e-118;
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGGACGGCCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Db 1 ATGCTACGAACCGCAGGACGGCCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Qy 61 GAGCGTGTGGAACGAAAGTTCAAGTTATACCTGGGACCGGACAGAGCTGGAGAA 120
 Db 61 GAGCGTGTGGAACGAAAGTTCAAGTTATACCTGGGACCGGACAGAGCTGGAGAA 120
 Qy 121 GCGAAGATCCCTCGGGAAGCATGGAAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Db 121 GCGAAGATCCCTCGGGAAGCATGGAAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Qy 181 ATCACCCTTGGGCGCAGAGAGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 240
 Db 181 ATCACCCTTGGGCGCAGAGAGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 240
 Qy 241 AACAGGAGGACCTTGGGAGAGGACAGAGAGGACCTGGTGAGGG 282
 Db 241 AACAGGAGGACCTTGGGAGAGGACAGAGAGGACCTGGTGAGGG 282

RESULT 15

AAK568757
 ID AAK568757 standard; cDNA; 506 BP.

XX AC AAK568757;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #4561.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG04570.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 4561; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 506 BP; 108 A; 149 C; 148 G; 101 T; 0 U; 0 Other;
Query Match 8.2%; Score 254; DB 5; Length 506;
Best Local Similarity 100.0%; Pred. No. 3.1e-105; Indels 0; Gaps 0;
Matches 254; Conservative 0; Mismatches 0;
Qy 1614 ACTCTGAACGAGGAGACCCAGGAGCCACCTGGAGAGAGTCTCTGTGGAAGGTCTCGCC 1673
Db 33 ACTCTGAACGAGGAGACCCAGGAGCCACCTGGAGAGAGTCTCTGTGGAAGGTCTCGCC 92
Qy 1674 GCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCAC 1733
Db 93 GCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCAC 152
Qy 1734 CCTGCAGCAGGGCTCCTTGGAGTTCCTCAGCTGCTTACAGATCCAGGAGGAGGTT 1793
Db 153 CCTGCAGCAGGGCTCCTTGGAGTTCCTCAGCTGCTTACAGATCCAGGAGGAGGTT 212
Qy 1794 TATCCAGCAGGGCCTCAGGCCACTTCCAGGTGATCGTGGTTCAGCAACATTCCTCCAAAGT 1853
Db 213 TATCCAGCAGGGCCTCAGGCCACTTCCAGGTGATCGTGGTTCAGCAACATTCCTCCAAAGT 272
Qy 1854 GGAGCACATGGTCT 1867
Db 273 GGAGCACATGGTCT 286

Search completed: August 8, 2004, 10:45:49
Job time : 1135 secs